

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 16:38:22 ; Search time 42 Seconds

(without alignments)
1652.765 Million cell updates/sec

Title: US-10-049-742-11

Perfect score: 1433

Sequence: 1 MAGVPEDELNPFHVLGVEAT.....VPKGEAKPKRKKVRRPQR 269

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1433	100.0	412	4	Q96N59 homo sapien
2	1367	95.4	699	6	Q95J56 bos taurus
3	1365	95.3	701	11	Q925G7 rattus norv
4	1361	95.0	659	6	Q9BGH5 bos taurus
5	1361	95.0	703	11	Q9CYB7 mus musculu
6	1357	94.7	703	11	Q921R4 mus musculu
7	1354	94.5	703	11	Q8BTU3 mus musculu
8	1155	80.7	678	11	Q8BLF3 mus musculu
9	1143	79.8	213	4	Q96AM4 homo sapien
10	1027	71.7	389	4	Q9C062 homo sapien
11	841.5	58.7	191	4	Q9C062 homo sapien
12	775	54.1	417	12	Q9E2W0 bovine vira
13	766	53.5	419	12	Q9E2W2 bovine vira
14	688	48.0	309	12	Q65451 bovine vira
15	685	47.8	334	12	Q9E2W7 bovine vira
16	685	47.8	411	12	Q9E2W6 bovine vira

17	618.5	43.2	250	12	Q9E2W1 bovine vira
18	585.5	40.9	291	12	Q65450 bovine vira
19	571.5	39.9	560	11	Q8BKF3 mus musculu
20	571	39.8	409	12	Q9E2V8 bovine vira
21	565	39.4	239	12	Q65797 mucosal dis
22	561	39.1	283	12	Q9E2V9 bovine vira
23	549	38.3	237	12	Q65798 mucosal dis
24	534	37.3	3989	12	Q9PYB2 pestivirus
25	497	34.7	3988	12	Q91WA7 mucosal dis
26	497	34.7	4040	12	Q91FH8 mucosal dis
27	493	34.4	381	12	Q65787 bovine vira
28	489.5	34.2	371	12	Q9E2W3 bovine vira
29	416	29.0	648	5	Q961F2 drosophila
30	416	29.0	970	5	Q9VN28 drosophila
31	284.5	19.9	387	10	Q9FX81 arabidopsis
32	282.5	19.7	868	5	Q8MQF1 caenorhabdi
33	272.5	19.0	402	10	Q9ST16 brassica ca
34	272.5	19.0	695	10	Q9FGY8 arabidopsis
35	264.5	18.5	577	10	Q64346 arabidopsis
36	193	13.5	376	11	Q8K037 mus musculu
37	193	13.5	378	11	Q8C4C9 mus musculu
38	182	12.7	370	5	Q9VFP0 drosophila
39	179	12.5	379	4	Q8TBM8 homo sapien
40	172	12.0	744	10	Q8RYF9 oryza sativ
41	169	11.8	539	10	Q49475 arabidopsis
42	168.5	11.8	378	5	Q22028 caenorhabdi
43	167	11.7	656	10	Q9SLA7 arabidopsis
44	165.5	11.5	223	11	Q9QZW8 rattus norv
45	165.5	11.5	748	10	Q94H83 oryza sativ

ALIGNMENTS

RESULT 1

Q96N59 PRELIMINARY; PRT; 412 AA.
ID Q96N59
AC Q96N59
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ31383.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN NCBJ_TaxID=9606;
RX [1]
RP SEQUENCE FROM N.A.
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukushima Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Maehara Y., Nagai K.,
RA Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AK055945; BAB71050.1; --
DR InterPro; IPR001823; DnaJ_N.
DR InterPro; IPR003095; Hsp_DnaJ.
DR Pfam; PF00226; DnaJ; 1.
DR PRINTS; PR00625; DnaJPROTEIN.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS50076; DnaJ_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 412 AA; 46931 MW; B51386515456C378 CRC64;

Query Match 100.0%; Score 1433; DB 4; Length 412;
Best Local Similarity 100.0%; Pred. No. 5,4e-115;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGVPEDELNPFHVLGVEATSDVELKKAYRQLAVMVHPDKXHHPRAEAFKVLRAWDI 60

```
Db 144 MAGVPEDELNPFHVLGVATASDELKAYROLAVMVHPDKNHHPRABEAFKVLRAAWDI 203
Qy 61 VSNAEKREYEMKMAENELSRVNEFLSKLQDDLKEAMNTMWCRCQCKHRRFEMDREP 120
Db 204 VSNAEKREYEMKMAENELSRVNEFLSKLQDDLKEAMNTMWCRCQCKHRRFEMDREP 263
Qy 121 KSARYCAECNRLHPAEEDGFWAESMLGLKITFYFALMDGKVYDITWAGCQRVGISPDTH 180
Db 264 KSARYCAECNRLHPAEEDGFWAESMLGLKITFYFALMDGKVYDITWAGCQRVGISPDTH 323
Qy 181 RVPYHISFGSRIPGTRGRATPDAPPADLQDPLSLRIFQVPPQMGNGNFFFAAPQAPGA 240
Db 324 RVPYHISFGSRIPGTRGRATPDAPPADLQDPLSLRIFQVPPQMGNGNFFFAAPQAPGA 383
Qy 241 AAASKPNSTVPKGEAKPKRRKKVRRPQR 269
Db 384 AAASKPNSTVPKGEAKPKRRKKVRRPQR 412
Qy 241 AAASKPNSTVPKGEAKPKRRKKVRRPQR 269
Db 384 AAASKPNSTVPKGEAKPKRRKKVRRPQR 412
PRT; 699 AA.
RESULT 2
Q95J56 PRELIMINARY;
AC Q95J56;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE J-domain protein Jiv.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
NCBI_TaxID=9913;
[1]
SEQUENCE FROM N.A.
MEDLINE=21424530; PubMed=11533209;
RA Rinck G., Birghan C., Harada T., Meyers G., Thiel H.J., Tautz N.;
RT "A cellular J-domain protein modulates polyprotein processing and
RT cytopathogenicity of a pestivirus.";
RL J. Virol. 75:9470-9482(2001).
DR EMEL; AY027882; AAK28651.1; -.
DR ENBL; AY027881; AAK28650.1; -.
DR InterPro; IPR001623; DnaJ_N.
DR Pfam; PF00226; DnaJ_1.
DR PRINTS; PR00625; DNAJPROTEIN.
DR SMART; SM00271; DnaJ_1.
DR PROSITE; PS50076; DnaJ_2; 1.
DR PROSITE; PS50076; DnaJ_2; 1.
DR Receptor.
KW Receptor.
SQ SEQUENCE 699 AA; 78178 MW; 370F62297595F516 CRC64;
Query Match 95.4%; Score 1367; DB 6; Length 699;
Best Local Similarity 95.9%; Pred. NO. 5e-109;
Matches 258; Conservative 2; Mismatches 5; Indels 4; Gaps 1;
Qy 1 MAGVPEDELNPFHVLGVATASDELKAYROLAVMVHPDKNHHPRABEAFKVLRAAWDI 60
Db 435 MAGVPEDELNPFHVLGVATASDELKAYROLAVMVHPDKNHHPRABEAFKVLRAAWDI 494
Qy 61 VSNAEKREYEMKMAENELSRVNEFLSKLQDDLKEAMNTMWCRCQCKHRRFEMDREP 120
Db 495 VSNPERRKEYEMKMAENELSRVNEFLSKLQ-----EAMNTMWCRCQCKHRRFEMDREP 550
Qy 121 KSARYCAECNRLHPAEEDGFWAESMLGLKITFYFALMDGKVYDITWAGCQRVGISPDTH 180
Db 551 KSARYCAECNRLHPAEEDGFWAESMLGLKITFYFALMDGKVYDITWAGCQRVGISPDTH 610
Qy 181 RVPYHISFGSRIPGTRGRATPDAPPADLQDPLSLRIFQVPPQMGNGNFFFAAPQAPGA 240
Db 611 RVPYHISFGSRIPGTRGRATPDAPPADLQDPLSLRIFQVPPQMGNGNFFFAAPQAPGA 670
Qy 241 AAASKPNSTVPKGEAKPKRRKKVRRPQR 269
Db 671 AAASKPNSTVPKGEAKPKRRKKVRRPQR 699
```

```
RESULT 3
Q925G7 PRELIMINARY; PRT; 701 AA.
AC Q925G7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Dopamine receptor interacting protein.
GN DRIP78.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=101116;
[1]
SEQUENCE FROM N.A.
STRAIN=Sprague-Dawley;
MEDLINE=21213175; PubMed=11331877;
RA Bermak J.C., Li M., Bullock C.M., Zhou Q.-Y.;
RT "Regulation of transport of the dopamine D1 receptor by a new
RT membrane-associated ER protein.";
RL Nat. Cell Biol. 3:492-498(2001).
DR EMEL; AF351783; AAK56240.1; -.
DR InterPro; IPR001623; DnaJ_N.
DR InterPro; IPR003095; Hsp_DnaJ.
DR Pfam; PF00226; DnaJ_1.
DR PRINTS; PR00625; DNAJPROTEIN.
DR SMART; SM00271; DnaJ_1.
DR PROSITE; PS50076; DnaJ_2; 1.
DR PROSITE; PS50076; DnaJ_2; 1.
DR Receptor.
KW Receptor.
SQ SEQUENCE 701 AA; 78578 MW; C8DCFC7BAC79BF19 CRC64;
Query Match 95.3%; Score 1365; DB 11; Length 701;
Best Local Similarity 93.7%; Pred. NO. 7.4e-109;
Matches 252; Conservative 8; Mismatches 9; Indels 0; Gaps 0;
Qy 1 MAGVPEDELNPFHVLGVATASDELKAYROLAVMVHPDKNHHPRABEAFKVLRAAWDI 60
Db 433 MAGVPEDELNPFHVLGVATASDELKAYROLAVMVHPDKNHHPRABEAFKVLRAAWDI 492
Qy 61 VSNAEKREYEMKMAENELSRVNEFLSKLQDDLKEAMNTMWCRCQCKHRRFEMDREP 120
Db 493 VSNPERRKEYEMKMAENELSRVNEFLSKLQDDLKEAMNTMWCRCQCKHRRFEMDREP 552
Qy 121 KSARYCAECNRLHPAEEDGFWAESMLGLKITFYFALMDGKVYDITWAGCQRVGISPDTH 180
Db 553 KSARYCAECNRLHPAEEDGFWAESMLGLKITFYFALMDGKVYDITWAGCQRVGISPDTH 612
Qy 181 RVPYHISFGSRIPGTRGRATPDAPPADLQDPLSLRIFQVPPQMGNGNFFFAAPQAPGA 240
Db 613 RVPYHISFGSRIPGTRGRATPDAPPADLQDPLSLRIFQVPPQMGNGNFFFAAPQAPGA 672
Qy 241 AAASKPNSTVPKGEAKPKRRKKVRRPQR 269
Db 673 TSTSRNPSTVPKGEAKPKRRKKVRRPQR 701
PRT; 659 AA.
RESULT 4
Q9BGH5 PRELIMINARY; PRT; 659 AA.
AC Q9BGH5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE DnaJ protein (fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
NCBI_TaxID=9913;
[1]
SEQUENCE FROM N.A.
```

RA Neill J.D., Ridpath J.F.;
 RT "Recombination with a cellular mRNA encoding a novel DnaJ protein."
 RT results in biotype conversion in genotype 2 bovine viral diarrhoea
 RT viruses";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF308815; AAG59810.2; -;
 DR HSSP; P25685; 1HDJ.
 DR InterPro; IPR001623; DnaJ_N.
 DR Pfam; PF00226; DnaJ_1.
 DR PRINTS; PR00625; DnaJPROTEIN.
 DR SMART; SM00271; DnaJ_1.
 DR PROSITE; PS50076; DnaJ_2; 1.
 FT NON TER 1
 SQ SEQUENCE 659 AA; 74130 MW; 9C576DDSD5E95839A CRC64;
 Query Match 95.0%; Score 1361; DB 6; Length 659;
 Best Local Similarity 95.5%; Pred. No. 1.5e-108;
 Matches 257; Conservative 2; Mismatches 6; Indels 4; Gaps 1;
 QY 1 MAGVPEDELNPFHVLGVEATASDELKKAQRLAVMVHPDKNHHPRAEAFKVLRAAWDI 60
 Db 395 MAGVPEDELNPFHVLGVEATASDELKKAQRLAVMVHPDKNHHPRAEAFKVLRAAWDI 454
 QY 61 VSNAEKREKEYEMKMAENELSRVNEFLSKLQDDLEAMNTMCSRCCGKHRRFEMDREP 120
 Db 455 VSNERRKEYEMKMAENELSRVNEFLSKLQDDLEAMNTMCSRCCGKHRRFEMDREP 510
 QY 121 KSARYCAECNRLHPAEEDGFWAESMLGLKITYPALMDGKVDITWAGCQVGLSPDTH 180
 Db 511 KSARYCAECNRLHPAEEDGFWAESMLGLKITYPALMDGKVDITWAGCQVGLSPDTH 570
 QY 181 RVPYHISFGSRIPGTRGQRATPDAPPADLQDFLSRIFQVPPGMPNGNFFAPOPAPGA 240
 Db 571 RVPYHISFGSRIPGTRGQRATPDAPPADLQDFLSRIFQVPPGMPNGNFFAPOPAPGA 630
 QY 241 AAASKPNSTVPKGEAKPKRRKKVRRPQR 269
 Db 631 TAASKPNSTVPKGEAKPKRRKKVRRPQR 659
 RESULT 5
 Q9CYB7 PRELIMINARY; PRT; 703 AA.
 ID Q9CYB7
 AC Q9CYB7; (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE 5730551F12Rik protein (Dopamine receptor interacting protein
 DE homolog).
 GN 5730551F12RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Harai A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilmink L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Brain;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; AK017830; BAB30962.1; -;
 DR HSSP; AK054445; BAC32372.1; -;
 DR HSSP; P25685; 1HDJ.
 DR MGD; MGI:1921580; 5730551F12Rik.
 DR InterPro; IPR001623; DnaJ_N.
 DR InterPro; IPR003095; Hsp_DnaJ.
 DR Pfam; PF00226; DnaJ_1.
 DR PRINTS; PR00625; DnaJPROTEIN.
 DR SMART; SM00271; DnaJ_1.
 DR PROSITE; PS50076; DnaJ_2; 1.
 DR SEQUENCE 703 AA; 78919 MW; B320DB4FB06345FE CRC64;
 SQ
 Query Match 95.0%; Score 1361; DB 11; Length 703;
 Best Local Similarity 93.3%; Pred. No. 1.6e-108;
 Matches 251; Conservative 8; Mismatches 10; Indels 0; Gaps 0;
 QY 1 MAGVPEDELNPFHVLGVEATASDELKKAQRLAVMVHPDKNHHPRAEAFKVLRAAWDI 60
 Db 435 MAGVPEDELNPFHVLGVEATASDELKKAQRLAVMVHPDKNHHPRAEAFKVLRAAWDI 494
 QY 61 VSNAEKREKEYEMKMAENELSRVNEFLSKLQDDLEAMNTMCSRCCGKHRRFEMDREP 120
 Db 495 VSNERRKEYEMKMAENELSRVNEFLSKLQDDLEAMNTMCSRCCGKHRRFEMDREP 554
 QY 121 KSARYCAECNRLHPAEEDGFWAESMLGLKITYPALMDGKVDITWAGCQVGLSPDTH 180
 Db 555 KSARYCAECNRLHPAEEDGFWAESMLGLKITYPALMDGKVDITWAGCQVGLSPDTH 614
 QY 181 RVPYHISFGSRIPGTRGQRATPDAPPADLQDFLSRIFQVPPGMPNGNFFAPOPAPGA 240
 Db 615 RVPYHISFGSRIPGTRGQRATPDAPPADLQDFLSRIFQVPPGMPNGNFFAPOPAPGA 674
 QY 241 AAASKPNSTVPKGEAKPKRRKKVRRPQR 269
 Db 675 TTSRPNSTVPKGEAKPKRRKKVRRPQR 703
 RESULT 6
 Q921R4 PRELIMINARY; PRT; 703 AA.
 ID Q921R4
 AC Q921R4;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE RIKEN cDNA 5730551F12 gene.
 GN 5730551F12RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 [1]
 RN SEQUENCE FROM N.A.
 RC Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC011146; AAH11146.1; -;
 DR MGD; MGI:1921580; 5730551F12Rik.
 DR InterPro; IPR001623; DnaJ_N.
 DR InterPro; IPR003095; Hsp_DnaJ.

DR Pfam: PF00226; DnaJ; 1.
DR PRINTS: PRO0625; DNAJPROTEIN.
DR SMART: SM00271; DnaJ; 1.
DR PROSITE: PS0076; DNAJ 2; 1.
SQ SEQUENCE 703 AA; 78931 MW; DC2ECB4FB64C47AB CRC64;

Query Match 94.7%; Score 1357; DB 11; Length 703;
Best Local Similarity 92.9%; Pred. No. 3.6e-108;
Matches 250; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 1 MAGVPEDELNPFHVLGVEATSDVELKAYRQLAVMVHPDKNHHPRAEAEAFKVLRAAWDI 60
DB 435 MAGVPEDELNPFHVLGVEATSDVELKAYRQLAVMVHPDKNHHPRAEAEAFKILRAAWDI 494
QY 61 VSNAEKREYEMKRMENELSRVNEFLSKLQDDLKEAMNTMCSRCCGKHRRFEMDREP 120
DB 495 VSNPERREYEMKRMENELSRVNEFLSKLQDDLKEAMNTMCSRCCGKHRRFEMDREP 554
QY 121 KSARYCAECNRLHPAEGDFWAESSMLGLKITTFALMDGKVDITWAGCQGVISPDTH 180
DB 555 KSARYCAECNRLHPAEGDFWAESSMLGLKITTFALMDGKVDITWAGCQGVISPDTH 614
QY 181 RVPYHISFGSRIPGTRGRQRTAPPADLQDFLSRIFQVPPGMPNGNFFAAPGP 240
DB 615 RVPYHISFGSRVPGTSGRQRATPESPVDLQDFLSRIFQVPPGMPNGNFFAAPHGP 674
QY 241 AAASKPNSTVPKGEAKPKRRKKVRRP 269
DB 675 TSTRPNSSVPGKEAKPKRRKKVRRP 703

Query Match 80.7%; Score 1156; DB 11; Length 678;
Best Local Similarity 90.8%; Pred. No. 6.7e-91;
Matches 216; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 1 MAGVPEDELNPFHVLGVEATSDVELKAYRQLAVMVHPDKNHHPRAEAEAFKVLRAAWDI 60
DB 435 MAGVPEDELNPFHVLGVEATSDVELKAYRQLAVMVHPDKNHHPRAEAEAFKILRAAWDI 494
QY 61 VSNAEKREYEMKRMENELSRVNEFLSKLQDDLKEAMNTMCSRCCGKHRRFEMDREP 120
DB 495 VSNPERREYEMKRMENELSRVNEFLSKLQDDLKEAMNTMCSRCCGKHRRFEMDREP 554
QY 121 KSARYCAECNRLHPAEGDFWAESSMLGLKITTFALMDGKVDITWAGCQGVISPDTH 180
DB 555 KSARYCAECNRLHPAEGDFWAESSMLGLKITTFALMDGKVDITWAGCQGVISPDTH 614
QY 181 RVPYHISFGSRIPGTRGRQRTAPPADLQDFLSRIFQVPPGMPNGNFFAAPGP 238
DB 615 RVPYHISFGSRVPGTSGRQRATPESPVDLQDFLSRIFQVPPGMPNGNFFAAPTSP 672

Query Match 94.5%; Score 1354; DB 11; Length 703;
Best Local Similarity 92.9%; Pred. No. 6.6e-108;
Matches 250; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 1 MAGVPEDELNPFHVLGVEATSDVELKAYRQLAVMVHPDKNHHPRAEAEAFKVLRAAWDI 60
DB 435 MAGVPEDELNPFHVLGVEATSDVELKAYRQLAVMVHPDKNHHPRAEAEAFKILRAAWDI 494
QY 61 VSNAEKREYEMKRMENELSRVNEFLSKLQDDLKEAMNTMCSRCCGKHRRFEMDREP 120
DB 495 VSNPERREYEMKRMENELSRVNEFLSKLQDDLKEAMNTMCSRCCGKHRRFEMDREP 554
QY 121 KSARYCAECNRLHPAEGDFWAESSMLGLKITTFALMDGKVDITWAGCQGVISPDTH 180
DB 555 KSARYCAECNRLHPAEGDFWAESSMLGLKITTFALMDGKVDITWAGCQGVISPDTH 614
QY 181 RVPYHISFGSRIPGTRGRQRTAPPADLQDFLSRIFQVPPGMPNGNFFAAPGP 240

Db 615 RVPYHISFGSRVPGTSGRQRATPESPVDLQDFLSRIFQVPPGMPNGNFFAAPHGP 674
QY 241 AAASKPNSTVPKGEAKPKRRKKVRRP 269
Db 675 TSTRPNSSVPGKEAKPKRRKKVRRP 703

RESULT 8
Q8BLF3 PRELIMINARY; PRT; 678 AA.
ID Q8BLF3
AC Q8BLF3
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Dopamine receptor interacting protein homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
DR EMBL; AK045345; BAC32320.1; -.
SQ SEQUENCE 678 AA; 76463 MW; 005FEFF7F9A5E5F CRC64;

Query Match 80.7%; Score 1156; DB 11; Length 678;
Best Local Similarity 90.8%; Pred. No. 6.7e-91;
Matches 216; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 1 MAGVPEDELNPFHVLGVEATSDVELKAYRQLAVMVHPDKNHHPRAEAEAFKVLRAAWDI 60
DB 435 MAGVPEDELNPFHVLGVEATSDVELKAYRQLAVMVHPDKNHHPRAEAEAFKILRAAWDI 494
QY 61 VSNAEKREYEMKRMENELSRVNEFLSKLQDDLKEAMNTMCSRCCGKHRRFEMDREP 120
DB 495 VSNPERREYEMKRMENELSRVNEFLSKLQDDLKEAMNTMCSRCCGKHRRFEMDREP 554
QY 121 KSARYCAECNRLHPAEGDFWAESSMLGLKITTFALMDGKVDITWAGCQGVISPDTH 180
DB 555 KSARYCAECNRLHPAEGDFWAESSMLGLKITTFALMDGKVDITWAGCQGVISPDTH 614
QY 181 RVPYHISFGSRIPGTRGRQRTAPPADLQDFLSRIFQVPPGMPNGNFFAAPGP 238
DB 615 RVPYHISFGSRVPGTSGRQRATPESPVDLQDFLSRIFQVPPGMPNGNFFAAPTSP 672

RESULT 9
Q96T63 PRELIMINARY; PRT; 213 AA.
ID Q96T63
AC Q96T63
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Dopamine receptor interacting protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21231375; PubMed=11331877;
RA Bermak J.C., Li M., Bullock C.M., Zhou Q.-Y.;
RT "Regulation of transport of the dopamine D1 receptor by a new
RT membrane-associated ER protein";
RL Nat. Cell Biol. 3:492-498(2001).
DR EMBL; AF351784; AAK36241.1; -.
KW Receptor.

```
FT NON TER 1 1
SQ SEQUENCE 213 AA; 24206 MW; E7AF40EAD9086613 CRC64;

Query Match
Best Local Similarity 79.8%; Score 1143; DB 4; Length 213;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 AWDIVSNAEKREKEYMKRMAENELSRVNEFLSKLQDDLEAMNTMCSRQCGKRRFEM 116
DB 1 AWDIVSNAEKREKEYMKRMAENELSRVNEFLSKLQDDLEAMNTMCSRQCGKRRFEM 60

QY 117 DREPKSARYCAECNRLHPAEGDFWAESSMLGLKITVFPALMDGKVDITWAGQCVGIS 176
DB 61 DREPKSARYCAECNRLHPAEGDFWAESSMLGLKITVFPALMDGKVDITWAGQCVGIS 120

QY 177 PDTHRVPHYHISFGSRIPGTRGRQATPDAPPADLQDFLSRIFQVPPGMPNGNFFAPOP 236
DB 121 PDTHRVPHYHISFGSRIPGTRGRQATPDAPPADLQDFLSRIFQVPPGMPNGNFFAPOP 180

QY 237 APGAAAASKPNSTVPKGEAKPKRRKVRFPQR 269
DB 181 APGAAAASKPNSTVPKGEAKPKRRKVRFPQR 213

RESULT 10
Q96AM4 PRELIMINARY; PRT; 389 AA.
AC Q96AM4
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC016941; AAH16941.1; -.
DR InterPro; IPR003034; SAP.
DR Pfam; PF02037; SAP; 1.
DR SMART; SM00513; SAP; 1.
KW Hypothetical protein.
SQ SEQUENCE 389 AA; 43745 MW; 3F8D3C3868031E66 CRC64;

Query Match
Best Local Similarity 71.7%; Score 1027; DB 4; Length 389;
Matches 191; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 72 MKRMAENELSRVNEFLSKLQDDLEAMNTMCSRQCGKRRFEMDRPKSARYCAECNR 131
DB 1 MKRMAENELSRVNEFLSKLQDDLEAMNTMCSRQCGKRRFEMDRPKSARYCAECNR 60

QY 132 LHPAEGDFWAESSMLGLKITVFPALMDGKVDITWAGQCVGISPDTHRVPHYHISGSR 191
DB 61 LHPAEGDFWAESSMLGLKITVFPALMDGKVDITWAGQCVGISPDTHRVPHYHISGSR 120

QY 192 IPGTRGRQATPDAPPADLQDFLSRIFQVPPGMPNGNFFAPOP 251
DB 121 IPGTRGRQATPDAPPADLQDFLSRIFQVPPGMPNGNFFAPOP 180

QY 252 KGEAKPKRRKKV 263
DB 181 KGEAKPKRRKKL 192

RESULT 11
Q9C062 PRELIMINARY; PRT; 191 AA.
AC Q9C062;

Query Match
Best Local Similarity 54.1%; Score 775; DB 12; Length 417;
Matches 142; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 101 TMCSRQCGKRRFEMDRPKSARYCAECNRLHPAEGDFWAESSMLGLKITVFPALMDGK 160
```

```
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE LYST-interacting protein LIP6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tchernev V.T., Mansfield T.A., Giot L., Kumar A.M., Nandabalan K.,
RA Li Y., Mishra V.S., Dettler J.C., Rothberg J.M., Wallace M.R.,
RA Southwick F.S., Kingmore S.F.;
RT "Interactions of the Chediak-Higashi lysosomal-traffic regulator
RT protein with SNARE complex and signal transduction proteins.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF141342; AAG49445.1; -.
SQ SEQUENCE 191 AA; 21675 MW; C9AE7DAE2775B101 CRC64;

Query Match
Best Local Similarity 58.7%; Score 841.5; DB 4; Length 191;
Matches 163; Conservative 1; Mismatches 12; Indels 9; Gaps 3;

QY 72 MKRMAENELSRVNEFLSKLQDDLEAMNTMCSRQCGKRRFEMDRPKSARYCAECNR 131
DB 1 MKRMAENELSRVNEFLSKLQDDLEAMNTMCSRQCGKRRFEMDRPKSARYCAECNR 60

QY 132 LHPAEGDFWAESSMLGLKITVFPALMDGKVDITWAGQCVGISPDTHRVPHYHISGSR 191
DB 61 LHPAEGDFWAESSMLGLKITVFPALMDGKVDITWAGQCVGISPDTHRVPHYHISGSR 120

QY 192 IPGTRGRQATPDAPPADLQDFLSRIFQVPPGMPNGNFFAPOP 242
DB 121 IPGTRGRQATPDAPPADLQDFLSRIFQVPPGMPNGNFFAPOP 180

QY 243 ASKPN 247
DB 181 REKPN 185

RESULT 12
Q9E2W0 PRELIMINARY; PRT; 417 AA.
AC Q9E2W0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Polypeptide (Fragment).
OS Bovine viral diarrhoea virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Pestivirus.
OX NCBI_TaxID=136447;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BVDV2-SD1630C;
RX MEDLINE=20411455; PubMed=10954582;
RA Ridpath J.F., Neill J.D.;
RT "Detection and characterization of genetic recombination in cytopathic
RT type 2 bovine viral diarrhoea viruses.";
RL J. Virol. 74:8771-8774 (2000).
DR EMBL; AF268178; AAG13371.1; -.
DR MEROPS; S31.001; -.
DR InterPro; IPR000280; CDvir_endptsep80.
DR PRINTS; PR00729; CDVENDOPTASE.
FT NON_TER 1
FT NON_TER 417
SQ SEQUENCE 417 AA; 46450 MW; FCDFF5F08DE58B9F CRC64;

Query Match
Best Local Similarity 92.8%; Score 775; DB 12; Length 417;
Matches 142; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 101 TMCSRQCGKRRFEMDRPKSARYCAECNRLHPAEGDFWAESSMLGLKITVFPALMDGK 160
```

```

82  TMWCSRCQGHRRLEMDREPKSARYCAECNRLHPAEEGDFWAESSMGLKITYFALMDGK 141
161  VYDITWAGCORVIGISPDTHRPVPHISFGSRIPGTRGRQRATPDAPPADLQDFLSRIFQV 220
142  VYDITWAGCORVIGISPDTHRPVPHISFGSRIPGTRGRQRATPDAPPADLQDFLSRIFQV 201
221  PGQMPNGNFFAAPQAPGAAASKPNSTVPKG 253
202  PGQMSNGNFFAAPQAPGATAASKPNREYKG 234

RESULT 13
Q9E2W2 PRELIMINARY; PRT; 419 AA.
ID Q9E2W2
AC Q9E2W2; 2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Polyprotein (Fragment).
OS Bovine viral diarrhea virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Pestivirus.
CX NCBI_TaxID=136447;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BVDV2-Galena 16425C;
RX MEDLINE=20411455; PubMed=10954582;
RA Ridpath J.F., Neill J.D.;
RT "Detection and characterization of genetic recombination in cytopathic
RT type 2 bovine viral diarrhea viruses.";
RL J. Virol. 74:8771-8774(2000).
DR EMBL; AF268176; AAG13369.1; -.
DR MEROPS; S31.001; -.
DR InterPro; IPR000280; CDvir_endptsep80.
DR PRINTS; PR00729; CDVENDOPTASE.
FT NON_TER 1
FT NON_TER 419
SQ SEQUENCE 419 AA; 46118 MW; 7E4028A0E018195D CRC64;

Query Match 53.5%; Score 766; DB 12; Length 419;
Best Local Similarity 93.4%; Pred. No. 1.3e-57;
Matches 141; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 95 LKEAMNTMCSRCCQGHRRLEMDREPKSARYCAECNRLHPAEEGDFWAESSMGLKITYF 154
DB 83 LKEAMITMCSRCCQGHRRLEMDREPKSARYCAECNRLHPAEEGDFWAESSMGLKITYF 142
QY 155 ALMDGKVYDITWAGCORVIGISPDTHRPVPHISFGSRIPGTRGRQRATPDAPPADLQDFL 214
DB 143 ALMDGKVYDITWAGCORVIGISPDTHRPVPHISFGSRIPGTRGRQRATPDAPPADLQDFL 202
QY 215 SRIFQVPPQMPNGNFFAAPQAPGAAASK 245
DB 203 SRIFQVPPQMSNGNFFAAPQPGGATAASKR 233

RESULT 14
Q65451 PRELIMINARY; PRT; 309 AA.
ID Q65451
AC Q65451; 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Nonstructural protein NS2-3 (Fragment).
GN NS2-3.
OS Bovine viral diarrhea virus genotype 2 (BVDV-2).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Pestivirus.
CX NCBI_TaxID=54315;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CP Cumnock;

```

```

RX MEDLINE=96186726; PubMed=8627775;
RA Becher P., Meyers G., Shannon A.D., Thiel H.J.;
RT "Cytopathogenicity of border disease virus is correlated with
RT integration of cellular sequences into the viral genome.";
RL J. Virol. 70:2992-2998(1996).
DR EMBL; U43603; AAB19180.1; -.
FT NON_TER 1
FT NON_TER 309
SQ SEQUENCE 309 AA; 35110 MW; 3A11237D6572792D CRC64;

Query Match 48.0%; Score 688; DB 12; Length 309;
Best Local Similarity 92.8%; Pred. No. 4.4e-51;
Matches 129; Conservative 1; Mismatches 5; Indels 4; Gaps 1;

QY 74 RMAENELSRVNEFLSKLQDDLKEAMNTMCSRCCQGHRRLEMDREPKSARYCAECNRLH 133
DB 64 REAENELSRVNEFLSKLQ----EAMNTMCSRCCQGHRRLEMDREPKSARYCAECNRLH 119
QY 134 PAEEGDFWAESSMGLKITYFALMDGKVYDITWAGCORVIGISPDTHRPVPHISFGSRIP 193
DB 120 PAEEGDFWAESSMGLKITYFALMDGKVYDITWAGCORVIGISPDTHRPVPHISFGSRIP 179
QY 194 GTRGRQRATPDAPPADLQD 212
DB 180 GTSGRQRATPDAPPADQSD 198

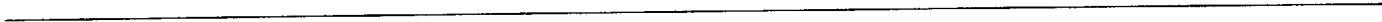
RESULT 15
Q9E2W7 PRELIMINARY; PRT; 334 AA.
ID Q9E2W7
AC Q9E2W7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Polyprotein (Fragment).
OS Bovine viral diarrhea virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Pestivirus.
CX NCBI_TaxID=136447;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BVDV2-296nc;
RX MEDLINE=20411455; PubMed=10954582;
RA Ridpath J.F., Neill J.D.;
RT "Detection and characterization of genetic recombination in cytopathic
RT type 2 bovine viral diarrhea viruses.";
RL J. Virol. 74:8771-8774(2000).
DR EMBL; AF268171; AAG13364.1; -.
DR MEROPS; S31.001; -.
DR InterPro; IPR000280; CDvir_endptsep80.
DR PRINTS; PR00729; CDVENDOPTASE.
FT NON_TER 1
FT NON_TER 334
SQ SEQUENCE 334 AA; F09D29D13D305476 CRC64;

Query Match 47.8%; Score 685; DB 12; Length 334;
Best Local Similarity 93.2%; Pred. No. 8.7e-51;
Matches 124; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 91 LQDDLKEAMNTMCSRCCQGHRRLEMDREPKSARYCAECNRLHPAEEGDFWAESSMGLK 150
DB 10 IREGCQEAAMNTMCSRCCQGHRRLEMDREPKSARYCAECNRLHPAEEGDFWAESSMGLK 69
QY 151 ITYFALMDGKVYDITWAGCORVIGISPDTHRPVPHISFGSRIPGTRGRQRATPDAPPADL 210
DB 70 ITYFALMDGKVYDITWAGCORVIGISPDTHRPVPHISFGSRIPGTRGRQRATPDAPPADL 129
QY 211 QDFLSRIFQVPPG 223
DB 130 QDFLSRIFQVPPG 142

Search completed: January 2, 2004, 16:42:28

```



GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 16:36:26 ; Search time 42 Seconds

(without alignments)
1016.606 Million cell updates/sec

Title: US-10-049-742-11

Perfect score: 1433

Sequence: 1 MAGVPEDELPHVLGVEAT.....VPKGEAKPKKVRPFQR 269

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :			
A_Geneseq_19Jun03.*			
1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*	1107863	
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*		
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*		
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*		
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*		
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*		
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*		
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*		
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*		
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*		
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*		
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*		
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*		
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*		
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*		
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*		
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*		
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*		
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*		
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*		
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*		
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*		
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*		
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1433	100.0	269	22	AAB67455 Amino acid sequenc
2	1430	99.8	412	23	ABG77019 Human protein sequ
3	1430	99.8	412	23	ABG64927 Human albumin fusi
4	1430	99.8	702	23	ABB77732 Amino acid sequenc
5	1367	95.4	699	23	ABB77490 Bovine Jiv protein
6	841.5	58.7	191	20	AA1932126 Human LYST interrac
7	841.5	58.7	191	23	ABG963164 Human ovarian canc
8	502	35.0	90	23	AA171492 Human Jiv protein
9	499	34.8	90	23	AA171491 Bovine Jiv protein

10	497	34.7	3835	22	AAB31167 Amino acid sequenc
11	497	34.7	3988	21	AA153615 NADL protein encod
12	497	34.7	3988	21	AA153615 Amino acid sequenc
13	416	29.0	970	22	ABG69643 Drosophila melanog
14	283	19.7	204	23	ABG77045 Human protein sequ
15	283	19.7	204	23	ABG64926 Human albumin fusi
16	189	13.2	258	22	AA171596 Novel signal trans
17	189	13.2	375	21	AA191941 Human chaperone pr
18	189	13.2	375	23	ABG7209 Human human protei
19	189	13.2	377	22	AA178619 Drosophila melanog
20	182	12.7	370	22	ABG59610 Novel signal trans
21	179	12.5	373	22	AA17224 Novel human diagno
22	177	12.4	408	22	ABG09740 Human ABCA1 intera
23	172.5	12.0	412	24	ABU11880 Novel signal trans
24	172	12.0	297	22	AA17597 Novel signal trans
25	171	11.9	397	24	ABJ18453 Breast specific re
26	170	11.9	276	22	AA19523 Human diagnostic a
27	165	11.5	216	22	AA17594 Novel signal trans
28	165	11.5	304	22	ABG3015 Human protein sequ
29	165	11.5	304	22	ABG7446 Amino acid sequenc
30	165	11.5	304	23	ABG7434 Novel human protei
31	165	11.5	312	22	AA17220 Novel signal trans
32	163	11.4	332	22	AAE01502 Human gene 21 enco
33	163	11.4	332	23	ABG3905 Human albumin fusi
34	163	11.4	415	20	AA107061 Renal cancer, assoc
35	163	11.4	737	22	AAE01456 Human gene 21 enco
36	163	11.4	737	23	ABG63906 Human albumin fusi
37	163	11.4	803	22	AA179322 Human protein SEQ
38	163	11.4	1052	22	AA178338 Human OREF ORF1748
39	162	11.3	337	21	AA141984 Human DnaJ protein
40	162	11.3	358	22	AA179625 Human protein sequ
41	162	11.3	358	22	AA195628 Amino acid sequenc
42	162	11.3	358	22	AA179628 Human polypeptide
43	162	11.3	397	22	AA138867 Human gene 21 enco
44	161	11.2	68	22	AAE01532 Novel signal trans
45	161	11.2	132	22	AA17598

ALIGNMENTS

RESULT 1

AAB67455

ID AAB67455 standard; Protein; 269 AA.

AC AAB67455;

XX 15-MAY-2001 (first entry)

DT Amino acid sequence of a human chaperone polypeptide.

DE Human; chaperone polypeptide; reproductive disease; prolactin production; infertilit; tumour; cancer; Peyronie's disease; eye disorder; glaucoma; conjunctivitis; keratitis; neuromuscular disorder; cystic fibrosis; metabolic disorder; Zellweger syndrome; Addison's disease; iritis; autoimmune disorder; inflammatory disorder; systemic lupus erythematosus; acquired immunodeficiency syndrome; AIDS; asthma; atherosclerosis; cell proliferative disorder; gene therapy.

OS Homo sapiens.

XX WO200109178-A2.

PN 08-FEB-2001.

PD 03-AUG-2000; 2000WO-US21313.

XX 03-AUG-1999; 99US-0146908.

PR 22-OCT-1999; 99US-0160924.

XX (INCY-) INCYTE GENOMICS INC.

PA Yue H, Bandman O, Tang YT, Baughn MR, Azimzai Y, Lu DAM;

PI

XX WPI: 2001-159853/16.
DR N-PSDB; AAF54994.
XX New human chaperone proteins and polynucleotides, useful in diagnosing,
PT treating and preventing reproductive, eye, neuromuscular, metabolic,
PT autoimmune or inflammatory disorders -
XX
XX Claim 1; Page 101-102; 102pp; English.
XX
XX The present sequence represents a human chaperone polypeptide. Human
CC chaperone polypeptides and polynucleotides are useful in the diagnosis,
CC treatment and prevention of reproductive (e.g. prolactin production,
CC infertility, endometrial or ovarian tumour, cancer of the breast,
CC prostate or testis, Peyronie's disease), eye (e.g. conjunctivitis,
CC keratitis, iritis, glaucoma), neuromuscular, metabolic (e.g. Zellweger
CC syndrome, Addison's disease, cystic fibrosis), and autoimmune and
CC inflammatory disorders (e.g. systemic lupus erythematosus, acquired
CC immunodeficiency syndrome (AIDS), asthma, atherosclerosis), infectious
CC or viral diseases, and cell proliferative disorders. Chaperone
CC polynucleotides may be used for somatic or germline gene therapy, to
CC detect and quantify gene expression in biopsied tissues in which
CC expression is correlated with disease.
XX
XX Sequence 269 AA;
SQ
Query Match 100.0%; Score 1433; DB 22; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.5e-132; Indels 0; Gaps 0;
Matches 269; Conservative 0; Mismatches 0;
QY 1 MAGVPEDELNPFHVLGVEATASDVVELKAYRQLAVMVHPDKNHHPRAEAFKVLRAAWDI 60
Db 1 MAGVPEDELNPFHVLGVEATASDVVELKAYRQLAVMVHPDKNHHPRAEAFKVLRAAWDI 60
QY 61 VSNAEKREKEYEMKMAENELSRVNEFLSKLQDDLKEAMNTMCSRQCGKHRRFEMDREP 120
Db 61 VSNAEKREKEYEMKMAENELSRVNEFLSKLQDDLKEAMNTMCSRQCGKHRRFEMDREP 120
QY 121 KSARYCAECNRLHPAEEDGFWAESMLGLKITYPALMDGKVDITWAGCQVGVISPDTH 180
Db 121 KSARYCAECNRLHPAEEDGFWAESMLGLKITYPALMDGKVDITWAGCQVGVISPDTH 180
QY 181 RVPYHISFGSRIPGTRGRQRATPDAPPADLQDLFLSRIFQVPPGQMPNGNFFAAPQAPGA 240
Db 181 RVPYHISFGSRIPGTRGRQRATPDAPPADLQDLFLSRIFQVPPGQMPNGNFFAAPQAPGA 240
QY 241 AAASKPNSTVPKGEAKPKERKKVRRPQR 269
Db 241 AAASKPNSTVPKGEAKPKERKKVRRPQR 269
RESULT 2
ABB77019
ID ABB77019 standard; Protein; 412 AA.
XX
AC ABB77019;
XX
XX 08-OCT-2002 (first entry)
DT
DE Human protein sequence #1 from clone HNTPB82.
XX
XX Human; HNTPB82; secreted protein; immunosuppressive; food preservative;
KW antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;
KW vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial;
KW virucide; fungicide; ophthalmological; vulnerary; gene therapy; ELISA;
KW radioimmunoassay; enzyme linked immunosorbent assay; autoimmune disease;
KW hyperproliferative disorder; cardiovascular disorder; angiogenesis;
KW cerebrovascular disorder; nervous system disorder; ocular disorder;
KW wound healing; food additive.
XX
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers

FT Peptide 1
FT Protein 2; 412
FT /label= Signal_peptide
FT /label= Mature_protein
XX WO200222638-A1.
XX
PD 21-MAR-2002.
XX
XX 17-JAN-2001; 2001WO-US01386.
XX
XX 12-SEP-2000; 2000US-232104P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;
PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;
PI Ni J;
XX WPI: 2002-258041/30.
XX N-PSDB; ABL55088.
XX
XX New nucleic acid molecules encoding 22 human secreted proteins for
PT diagnosing or treating e.g. autoimmune diseases, hyperproliferative
PT disorders, and cardiovascular disorders, and used as food additives or
PT preservatives -
XX
XX Disclosure: Page 486-487; 526pp; English.
XX
XX The sequence represents a protein sequence of the invention, encoded by
CC cDNA isolated from human clone ID HNTPB82. The invention relates to novel
CC isolated nucleic acid molecules encoding 22 human secreted proteins. The
CC proteins of the invention have immunosuppressive, antiarthritic,
CC antirheumatic, antiproliferative, cytostatic, cardiant, vasotropic,
CC cerebroprotective, neurotropic, neuroprotective, antibacterial, virucide,
CC fungicide, ophthalmological, and vulnerary activity. The polynucleotides
CC may have a use in gene therapy. The polynucleotides and polypeptides
CC encoded by them are used to prevent, treat or ameliorate a medical
CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
CC chickens or sheep. The polynucleotides and polypeptides are also used in
CC diagnosing a pathological condition or susceptibility to a pathological
CC condition. The antibodies to the proteins can also be used in alleviating
CC symptoms associated with the disorders and in diagnostic immunoassays
CC e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA).
CC Disorders which are diagnosed or treated include autoimmune diseases,
CC hyperproliferative disorders, cardiovascular disorders, infections caused by
CC disorders, angiogenesis, nervous system disorders, infections caused by
CC bacteria, viruses and fungi and ocular disorders. The polypeptides can
CC also be used to aid wound healing and epithelial cell proliferation. The
CC polypeptides can also be used as a food additive or preservative.
XX
SQ Sequence 412 AA;
Query Match 99.8%; Score 1430; DB 23; Length 412;
Best Local Similarity 99.8%; Pred. No. 5.3e-132; Indels 0; Gaps 0;
Matches 268; Conservative 1; Mismatches 0;
QY 1 MAGVPEDELNPFHVLGVEATASDVVELKAYRQLAVMVHPDKNHHPRAEAFKVLRAAWDI 60
Db 144 MAGVPEDELNPFHVLGVEATASDVVELKAYRQLAVMVHPDKNHHPRAEAFKVLRAAWDI 203
QY 61 VSNAEKREKEYEMKMAENELSRVNEFLSKLQDDLKEAMNTMCSRQCGKHRRFEMDREP 120
Db 204 VSNAEKREKEYEMKMAENELSRVNEFLSKLQDDLKEAMNTMCSRQCGKHRRFEMDREP 263
QY 121 KSARYCAECNRLHPAEEDGFWAESMLGLKITYPALMDGKVDITWAGCQVGVISPDTH 180
Db 264 KSARYCAECNRLHPAEEDGFWAESMLGLKITYPALMDGKVDITWAGCQVGVISPDTH 323
QY 181 RVPYHISFGSRIPGTRGRQRATPDAPPADLQDLFLSRIFQVPPGQMPNGNFFAAPQAPGA 240
Db 324 RVPYHISFGSRIPGTRGRQRATPDAPPADLQDLFLSRIFQVPPGQMPNGNFFAAPQAPGA 383

```

QY 241 AAASKPNSTVPKGEAKPKRRKKVRRPFR 269
DB 384 AAASKPNSTVPKGEAKPKRRKKVRRPFR 412

RESULT 3
ABG64927
ID ABG64927 standard; Protein; 412 AA.
XX
AC ABG64927;
DT 27-AUG-2002 (first entry)
DE Human albumin fusion protein #1602.
XX
XX Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW haematopoietic disorder; neural disorder; connective disorder;
KW cytostatic; antifertility; antiinflammatory; antiulcer;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nocotropic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.
XX
XX Homo sapiens.
OS Synthetic.
XX WO200177137-A1.
PN 18-OCT-2001.
XX 12-APR-2001; 2001WO-US11988.
XX 12-APR-2000; 2000US-229358P.
PR 25-APR-2000; 2000US-199384P.
PR 21-DEC-2000; 2000US-256931P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Haseltine WA;
PI WPI; 2002-010896/01.
XX
XX New fusion protein for treating disease e.g. diabetes comprises an
PT albumin fused to a therapeutic protein -
XX
XX Claim 1; Page 1606-1607; 2102pp; English.
XX
XX The present invention relates to albumin fusion proteins comprising a
CC therapeutic protein X and human albumin (HA, also known as human serum
CC albumin, HSA). The proteins are useful for treating a disease or
CC disorder that may be modulated by therapeutic protein X. The albumin
CC extends the shelf-life of protein X, and may increase its biological
CC in vitro/in vivo activity. The protein is useful for treating and
CC diagnosing disorders such as cancer, reproductive disorders, digestive
CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
CC (e.g. diabetes), haematopoietic disorders, neural disorders
CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
CC encephalomyelitis, meningitis, schizophrenia), and connective disorders
CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin
CC fusion proteins of the invention.
XX
XX Sequence 412 AA;
SQ
Query Match 99.8%; Score 1430; DB 23; Length 412;
Best Local Similarity 99.6%; Pred. No. 5.3e-132;
Matches 268; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGVPEDELNPFHVLGVATASDELKAYRQLAVMVHPDKNHPRAEAFKVLRAAWDI 60
DB 144 MAGVPEDELNPFHVLGVATASDELKAYRQLAVMVHPDKNHPRAEAFKVLRAAWDI 203

us-10-049-742-11.rag
QY 61 VSNAEKREYEMKMAENELSRVSNEFLSKLODDLKEAMNTMCSRCCQGHRRFEMDREP 120
DB 204 VSNAEKREYEMKMAENELSRVSNEFLSKLODDLKEAMNTMCSRCCQGHRRFEMDREP 263
QY 121 KSARYCAECNRLHPAEEGDFWAESSMLGLKITYFALMDGKVDITWAGCCQVIGISPDTH 180
DB 264 KSARYCAECNRLHPAEEGDFWAESSMLGLKITYFALMDGKVDITWAGCCQVIGISPDTH 323
QY 181 RVPYHISFGSRIPGTRGRORATPDAPPADLODFLSRIQVPPGOMPNGNFFFAAPAPGA 240
DB 324 RVPYHISFGSRIPGTRGRORATPDAPPADLODFLSRIQVPPGOMPNGNFFFAAPAPGA 383
QY 241 AAASKPNSTVPKGEAKPKRRKKVRRPFR 269
DB 384 AAASKPNSTVPKGEAKPKRRKKVRRPFR 412

RESULT 4
ABB77732
ID ABB77732 standard; Protein; 702 AA.
XX
AC ABB77732;
XX 30-JUL-2002 (first entry)
DT
XX
XX Amino acid sequence of human P125-77.22 polypeptide.
DE Human; P125-77.22; mucosal disease; BVDV infection; gene therapy.
XX
XX Homo sapiens.
XX WO200226810-A1.
XX 04-APR-2002.
XX 10-SEP-2001; 2001WO-CN01354.
XX 12-SEP-2000; 2000CN-0125190.
XX
XX (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.
PA
XX Mao Y, Xie Y;
PI
XX
XX WPI; 2002-281315/32.
XX N-PSDB; ABL56700.
XX
XX Protein P125-77.22 and encoding polynucleotide, used in diagnosis and
PT treatment of human mucosal disease caused by BVDV infection -
XX
XX Claim 1; Page 29-30; 33pp; Chinese.
XX
XX The present sequence represents human P125-77.22 polypeptide. The
CC polypeptide and polynucleotide are used in diagnosis and treatment
CC of human mucosal disease caused by viral BVDV (undefined) infection.
CC The polynucleotide may also be used for gene therapy.
XX
XX Sequence 702 AA;
SQ
Query Match 99.8%; Score 1430; DB 23; Length 702;
Best Local Similarity 99.6%; Pred. No. 1.1e-131;
Matches 268; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGVPEDELNPFHVLGVATASDELKAYRQLAVMVHPDKNHPRAEAFKVLRAAWDI 60
DB 434 MAGVPEDELNPFHVLGVATASDELKAYRQLAVMVHPDKNHPRAEAFKVLRAAWDI 493
QY 61 VSNAEKREYEMKMAENELSRVSNEFLSKLODDLKEAMNTMCSRCCQGHRRFEMDREP 120
DB 494 VSNAEKREYEMKMAENELSRVSNEFLSKLODDLKEAMNTMCSRCCQGHRRFEMDREP 553
QY 121 KSARYCAECNRLHPAEEGDFWAESSMLGLKITYFALMDGKVDITWAGCCQVIGISPDTH 180
DB 554 KSARYCAECNRLHPAEEGDFWAESSMLGLKITYFALMDGKVDITWAGCCQVIGISPDTH 613

```

181 RVPYHISFGSRIPGTRGRORATPDAPPADLQDLSRIQVPPGQMPNGNFFAAPQAPGA 240
614 RVPYHISFGSRIPGTRGRORATPDAPPADLQDLSRIQVPPGQMPNGNFFAAPQAPGA 673

241 AAASKPNSTVPKGEAKPKRRKVRPFQ 269
674 AAASKPNSTVPKGEAKPKRRKVRPFQ 702

RESULT 5
ID AAB71490
ID AAB71490 standard; protein; 699 AA.
XX AAB71490;
XX
DT 28-NOV-2002 (first entry)
XX
DE Bovine Jiv protein.
XX
KW NS2; non-structural protein 2; inhibitor; competitive; allosteric; Jiv;
KW J-domain protein interacting with viral protein; NS2-3 viral protease;
KW virucide; hepatotropic; antiinflammatory; infection; bovine.
XX
OS Bos taurus.
XX
PN DE10112748-A1.
XX
PD 19-SEP-2002.
XX
PF 14-MAR-2001; 2001DE-1012748.
XX
PR 14-MAR-2001; 2001DE-1012748.
XX
PA (TRAN-) TRANSMIT GES TECHNOLOGIETRANSFER MBH.
XX
PI Tautz N, Thiel H, Birghan C;
XX
DR WPI; 2002-692596/75.
XX
PT New binding partners for hepatitis C virus non-structural protein 2,
PT useful for diagnosis, prevention and treatment of hepatitis C infection
PT
XX
PS Disclosure; Fig 2A; 14pp; German.
XX
XX This invention describes a novel binding partner for non-structural
CC protein 2 (NS2) of hepatitis C virus (HCV) that competitively or
CC allosterically inhibits binding of Jiv (J-domain protein interacting
CC with viral protein) to NS2. Inhibiting Jiv and NS2 binding prevents
CC stimulation of NS2-3 viral protease. The products of the invention have
CC virucide, hepatotropic and antiinflammatory activity. The novel
CC binding partner is used for diagnosis, prevention and treatment of HCV
CC infection. This sequence represents the bovine Jiv protein described in
CC the disclosure of the invention.
XX
SQ Sequence 699 AA;
Query Match 95.4%; Score 1367; DB 23; Length 699;
Best Local Similarity 95.9%; Pred. No. 1.8e-125;
Matches 258; Conservative 2; Mismatches 5; Indels 4; Gaps 1;

Qy 1 MAGVPEDELNPFHVLGVATASDVELKKAYROLAVMWHDPKHHFRAEEAFKVLRAAWDI 60
Db 435 MAGVPEDELNPFHVLGVATASDVELKKAYROLAVMWHDPKHHFRAEEAFKVLRAAWDI 494

Qy 61 VSNKREKREYEMKRAENELSSVNEFLSKLODDLKEAMNTWMCSCQKRRFEMDREP 120
Db 495 VSNPERRREYEMKRAENELSSVNEFLSKLODDLKEAMNTWMCSCQKRRFEMDREP 550

Qy 121 KSARYCAECNRLHPAEEGDFWAESSMLGLKITTYFALMDGKVYDITWAGCQRVGISPDTH 180
Db 551 KSARYCAECNRLHPAEEGDFWAESSMLGLKITTYFALMDGKVYDITWAGCQRVGISPDTH 610

Qy 181 RVPYHISFGSRIPGTRGRORATPDAPPADLQDLSRIQVPPGQMPNGNFFAAPQAPGA 240
Db 611 RVPYHISFGSRIPGTRGRORATPDAPPADLQDLSRIQVPPGQMPNGNFFAAPQAPGA 670

Qy 241 AAASKPNSTVPKGEAKPKRRKVRPFQ 269
Db 671 TAASKPNSTVPKGEAKPKRRKVRPFQ 699

RESULT 6
ID AAY32126
ID AAY32126 standard; Protein; 191 AA.
XX
AC AAY32126;
XX
DT 01-FEB-2000 (first entry)
XX
DE Human LYST interacting protein LIP6.
XX
KW LIP6; human; LYST-2; LYST interacting protein; lysosome; CHS;
KW Chediak-Higashi syndrome; autoimmune disease; rheumatoid arthritis;
KW systemic lupus erythematosus; inflammatory bowel disease;
KW diabetes mellitus; multiple sclerosis; atopic disease; asthma;
KW hay fever; rhinitis; urticaria; nasal polyp; cancer;
KW neurodegenerative disease; pigmentation disorder; viral disease;
KW platelet dysfunction.
XX
OS Homo sapiens.
XX
PN WO9951741-A2.
XX
PD 14-OCT-1999.
XX
PF 29-MAR-1999; 99WO-US06831.
XX
PR 03-APR-1998; 98US-0054956.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Nandabalan K, Kingmore S;
XX
DR WPI; 1999-620203/53.
DR N-PSDB; AAZ34492.
XX
PT Protein complexes, interacting proteins, and related polynucleotides
PT useful for treating and preventing e.g. atopic, autoimmune or
PT neurodegenerative diseases -
XX
PS Claim 20; Fig 7; 172pp; English.
XX
XX The present sequence represents a novel human LYST interacting
CC protein, LIP6, that shows homology to pestivirus type 3 NS2-3.
CC LYST is the human lysosomal Chediak-Higashi syndrome (CHS) protein.
CC The invention relates to complexes of LYST or LYST-2 (see AAY32126)
CC with proteins identified as interacting with LYST or LYST-2 by a
CC modified yeast two-hybrid assay system. The interacting proteins
CC include 10 novel proteins, LIP1-10 (see AAY32121-30). Methods of
CC screening the protein complexes for efficacy in treating and/or
CC preventing atopic diseases (e.g. asthma, nasal polyps, hay fever
CC rhinitis, urticaria) autoimmune diseases (e.g. CHS, rheumatoid
CC arthritis, systemic lupus erythematosus, inflammatory bowel disease,
CC diabetes mellitus, multiple sclerosis), neurodegenerative disease,
CC certain cancers, pigmentation disorders, platelet dysfunction and
CC viral diseases are provided. Nucleic acids (see AAZ34487-96)
CC encoding LIP1-10, modulation of LIP function by gene therapy, use
CC of antisense oligonucleotides for suppression of LIP protein
CC expression, screening for agonists and antagonists, diagnosing or
CC screening for the presence of a predisposition to a disease or
CC disorder, and animal models are also disclosed.
XX
SQ Sequence 191 AA;

Query Match 58.7%; Score 841.5; DB 20; Length 191;
Best Local Similarity 88.1%; Pred. No. 1.6e-74;
Matches 163; Conservative 1; Mismatches 12; Indels 9; Gaps 3;

QY 72 MKRMAENELSRVNEFLSKLQDDLEAMNTMCSRQCGKRRFEMDRPKSARYCAECNR 131
DB 1 MKRMAENELSRVNEFLSKLQDDLEAMNTMCSRQCGKRRFEMDRPKSARYCAECNR 60

QY 132 LHPAEEGDFWAESSMLGLKITYPALMDGKYDITEWAGCORVIGISPDTHRPVPHISFGSR 191
DB 61 LHPAEEGDFWAESSMLGLKITYPALMDGKYDITEWAGCORVIGISPDTHRPVPHISFGSR 120

QY 192 IPGTRGRQATPDAPPADLQDFLSRIFQVPPGOMPENG---NFFAAP---QPAPGAA--A 242
DB 121 IPGTRGRQATPDAPPADLQDFLSRIFQVPPGOMPENG---NFFAAP---QPAPGAA--A 242

QY 243 ASKPN 247
DB 181 REKPN 185

RESULT 7
ABG96364 standard; Protein; 191 AA.
XX
AC ABG96364;
DT 11-DEC-2002 (first entry)
XX
DE Human ovarian cancer marker OV38.
XX
KW Human, ovarian cancer; marker; cancer; familial history; brain disorder;
KW central nervous system disorder; bacterial meningitis; viral meningitis;
KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
KW brain herniation; inflammation; encephalitis; testicular disorder;
KW nontuberculous granulomatous orchitis; connective tissue disorder;
KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
KW histological type; carcinogenic; ovarian cancer marker.
XX
OS Homo sapiens.
XX
PN WO200271928-A2.
XX
PD 19-SEP-2002.
XX
PF 14-MAR-2002; 2002WO-US07826.
XX
PR 14-MAR-2001; 2001US-276025P.
PR 14-MAR-2001; 2001US-276026P.
PR 10-AUG-2001; 2001US-311732P.
PR 19-SEP-2001; 2001US-323580P.
PR 26-SEP-2001; 2001US-324967P.
PR 26-SEP-2001; 2001US-325102P.
PR 26-SEP-2001; 2001US-325149P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
PI Meyers RE, Morrisset MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
PI Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;
XX
XX WPI; 2002-723277/78.
DR N-PSDB; ABS76460.
XX
XX Assessing whether a patient is afflicted with ovarian cancer, useful in
PT assessing the stage or progression of the disease, comprises comparing
PT the expression level of a cancer marker in a sample from a patient and
PT from a non cancer patient -
XX
XX Disclosure; Page 300; 481pp; English.
PS
XX The present invention relates to a new method for assessing whether a
CC patient is afflicted with ovarian cancer. The method involves comparing

the expression level of a marker in a patient sample and the normal level
of expression of the marker in a control non-ovarian cancer sample, where
the marker is selected from 363 cancer markers described in the
specification. The method of the invention is useful in diagnosing or
characterising cancer, in detecting the presence of cancer as early as
possible, and the recurrence of ovarian cancer. The method may also be of
particular use with patients having an enhanced risk of developing
ovarian cancer (e.g. patients having a familial history of ovarian
cancer). The cancer markers may be used in the management and treatment
of e.g. brain and central nervous system disorders (e.g. bacterial and
viral meningitis, Alzheimer's disease or Parkinson's disease), brain
disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
inflammations (e.g. bacterial or viral meningitis or encephalitis),
testicular disorders (e.g. nontuberculous granulomatous orchitis),
connective tissue disorders, or heart disorders (e.g. ischaemic heart
disease or atherosclerosis). The compositions and methods may also be
used in assessing the histological type of neoplasm associated with
ovarian cancer, monitoring the progression of ovarian cancer,
determining whether ovarian cancer has metastasized or is likely to
metastasize, selecting a composition for inhibiting ovarian cancer,
assessing the ovarian carcinogenic potential of a compound, or
inhibiting ovarian cancer or at risk of developing ovarian cancer. The
present amino acid sequence represents one of the ovarian cancer markers
described in the invention.
SQ Sequence 191 AA;

Query Match 58.7%; Score 841.5; DB 23; Length 191;
Best Local Similarity 88.1%; Pred. No. 1.6e-74;
Matches 163; Conservative 1; Mismatches 12; Indels 9; Gaps 3;

QY 72 MKRMAENELSRVNEFLSKLQDDLEAMNTMCSRQCGKRRFEMDRPKSARYCAECNR 131
DB 1 MKRMAENELSRVNEFLSKLQDDLEAMNTMCSRQCGKRRFEMDRPKSARYCAECNR 60

QY 132 LHPAEEGDFWAESSMLGLKITYPALMDGKYDITEWAGCORVIGISPDTHRPVPHISFGSR 191
DB 61 LHPAEEGDFWAESSMLGLKITYPALMDGKYDITEWAGCORVIGISPDTHRPVPHISFGSR 120

QY 192 IPGTRGRQATPDAPPADLQDFLSRIFQVPPGOMPENG---NFFAAP---QPAPGAA--A 242
DB 121 IPGTRGRQATPDAPPADLQDFLSRIFQVPPGOMPENG---NFFAAP---QPAPGAA--A 242

QY 243 ASKPN 247
DB 181 REKPN 185

RESULT 8
AAB71492 standard; protein; 90 AA.
XX
AC AAB71492;
DT 28-NOV-2002 (first entry)
XX
DE Human Jiv protein Jiv90 fragment.
XX
KW NS2; non-structural protein 2; inhibitor; competitive; allosteric; Jiv;
KW J-domain protein interacting with viral protein; NS2-3 viral protease;
KW virucide; hepatotropic; antiinflammatory; infection; human; Jiv90.
XX
OS Homo sapiens.
XX
PN DE10112748-A1.
XX
PD 19-SEP-2002.
XX
PF 14-MAR-2001; 2001DE-1012748.
XX
PR 14-MAR-2001; 2001DE-1012748.
XX
PA (TRAN-) TRANSMIT GES TECHNOLOGIETRANSFER MBH.

XX Tautz N, Thiel H, Birghan C;
 XX WPI; 2002-692596/75.
 DR New binding partners for hepatitis C virus non-structural protein 2,
 PT useful for diagnosis, prevention and treatment of hepatitis C infection
 PT -
 XX Disclosure; Fig 3; 14pp; German.
 PS This invention describes a novel binding partner for non-structural
 CC protein 2 (NS2) of hepatitis C virus (HCV) that competitively or
 CC allosterically inhibits binding of Jiv (J-domain protein interacting
 CC with viral protein) to NS2. Inhibiting Jiv and NS2 binding prevents
 CC stimulation of NS2-3 viral protease. The products of the invention have
 CC virucide, hepatotropic and antiinflammatory activity. The novel
 CC binding partner is used for diagnosis, prevention and treatment of HCV
 CC infection. This sequence represents the human Jiv protein Jiv90 fragment
 CC described in the disclosure of the invention.
 XX
 SQ Sequence 90 AA;
 Query Match 35.0%; Score 502; DB 23; Length 90;
 Best Local Similarity 100.0%; Pred. No. 1.5e-41;
 Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 103 MCSRCQKRRRFEEMDRPKSARYCAECNRLHPAEGDFWAEESMLGLKITYFALMDGKY 162
 Db 1 MCSRCQKRRRFEEMDRPKSARYCAECNRLHPAEGDFWAEESMLGLKITYFALMDGKY 60
 QY 163 DITWAGCORVGISPDTHRPVPHISFGSRI 192
 Db 61 DITWAGCORVGISPDTHRPVPHISFGSRI 90
 RESULT 9
 AAB71491
 ID AAB71491 standard; protein; 90 AA.
 AC AAB71491;
 XX 28-NOV-2002 (first entry)
 DT Bovine Jiv protein Jiv90 fragment.
 DE NS2; non-structural protein 2; inhibitor; competitive; allosteric; Jiv;
 KW J-domain protein interacting with viral protein; NS2-3 viral protease;
 KW virucide; hepatotropic; antiinflammatory; infection; bovine; Jiv90.
 XX
 OS Bos taurus.
 XX DE10112748-A1.
 PN 19-SEP-2002.
 PD 14-MAR-2001; 2001DE-1012748.
 PF 14-MAR-2001; 2001DE-1012748.
 PR (TRAN-) TRANSMIT GES TECHNOLOGIETRANSFER MBH.
 PA Tautz N, Thiel H, Birghan C;
 XX WPI; 2002-692596/75.
 DR New binding partners for hepatitis C virus non-structural protein 2,
 XX useful for diagnosis, prevention and treatment of hepatitis C infection
 PT -
 PT Disclosure; Fig 28; 14pp; German.
 PS This invention describes a novel binding partner for non-structural
 CC

CC protein 2 (NS2) of hepatitis C virus (HCV) that competitively or
 CC allosterically inhibits binding of Jiv (J-domain protein interacting
 CC with viral protein) to NS2. Inhibiting Jiv and NS2 binding prevents
 CC stimulation of NS2-3 viral protease. The products of the invention have
 CC virucide, hepatotropic and antiinflammatory activity. The novel
 CC binding partner is used for diagnosis, prevention and treatment of HCV
 CC infection. This sequence represents the bovine Jiv protein Jiv90 fragment
 CC described in the disclosure of the invention.
 XX
 SQ Sequence 90 AA;
 Query Match 34.8%; Score 499; DB 23; Length 90;
 Best Local Similarity 98.9%; Pred. No. 3e-41;
 Matches 89; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 103 MCSRCQKRRRFEEMDRPKSARYCAECNRLHPAEGDFWAEESMLGLKITYFALMDGKY 162
 Db 1 MCSRCQKRRRFEEMDRPKSARYCAECNRLHPAEGDFWAEESMLGLKITYFALMDGKY 60
 QY 163 DITWAGCORVGISPDTHRPVPHISFGSRI 192
 Db 61 DITWAGCORVGISPDTHRPVPHISFGSRI 90
 RESULT 10
 AAB31167
 ID AAB31167 standard; protein; 3835 AA.
 XX AAB31167;
 AC 02-APR-2001 (first entry)
 DT Amino acid sequence of a chimeric BVDV/HCV virus.
 XX Chimeric virus; bovine viral diarrhoea virus; BVDV; hepatitis C virus;
 DE HCV; vaccine; viral inhibitor; antiviral.
 KW Synthetic.
 XX Bovine viral diarrhoea virus.
 OS Hepatitis C virus.
 OS WO200075352-A2.
 PN 14-DEC-2000.
 PD 02-JUN-2000; 2000WO-US15527.
 PF 04-JUN-1999; 99US-0137817.
 PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Nam J, Bukh J, Emerson SU, Purcell RH;
 XX WPI; 2001-071081/08.
 DR N-PSDB; AAC86936.
 XX New nucleic acid comprising a chimeric bovine viral diarrhoea virus
 PT genome in which the (non-)structural region has been replaced by
 PT hepatitis C virus (HCV) genome useful for treating or preventing HCV
 PT signs and symptoms -
 XX Disclosure; Page 66-81; 97pp; English.
 PS The specification describes a nucleic acid comprising a chimeric virus
 XX genome, specifically bovine viral diarrhoea virus (BVDV) genome in which
 CC the (non-)structural region has been replaced by the (non-)structural
 CC region of a hepatitis C virus (HCV) genome. The nucleic acids comprising
 CC the chimeric virus and the chimeric virus are useful for identifying
 CC cell lines capable of supporting the replication of these chimeric
 CC viruses, in screening for neutralizing antibodies to HCV of different
 CC genotypes, in the production of HCV-BVDV virions, for the development
 CC of inactivated or attenuated vaccines to prevent HCV-BVDV in a mammal,
 CC in studying the molecular properties of HCV indirectly in vitro, and in

CC identifying inhibitors of viral enzyme activity which would be useful
CC as antiviral agents. Formulations of compositions comprising the
CC chimeric virions may be used to treat or prevent the signs and symptoms
CC of HCV. The present sequence is encoded by a chimeric nucleic acid of the
CC invention.

XX SQ Sequence 3835 AA;
Query Match 34.7%; Score 497; DB 22; Length 3835;
Best Local Similarity 97.8%; Pred. No. 1.1e-38;
Matches 89; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 103 MCSRCQGHRRFEMDRPEKSAARYCAECNRLHPAEGDFWAESSMLGLKITYPALMDGKVY 162
DB 1384 MCSRCQGHRRFEMDRPEKSAARYCAECNRLHPAEGDFWAESSMLGLKITYPALMDGKVY 1443

QY 163 DITEWAGCQRVGISPDTHRVPHYHISFGSRIP 193
DB 1444 DITEWAGCQRVGISPDTHRVPHYHISFGSRMP 1474

RESULT 11
ID AAY53615 standard; Protein; 3988 AA.
XX AC AAY53615;
XX DT 11-FEB-2000 (first entry)
XX DE NADL protein encoded by the low copy number plasmid pACNR/BVDV NADL.
XX KW Pseudorevertant; RNA virus; chimera; BVDV; HCV; replication-competent;
XX KW 5' nontranslated region; 5'NTR; 3' NTR; pestivirus; antiviral;
XX KW bovine viral diarrhea virus; NADL; vaccine.
XX OS Synthetic.
XX OS Bovine viral diarrhea virus.
XX PN WO9955366-A1.
XX PD 04-NOV-1999.
XX PF 23-APR-1999; 99WO-US08850.
XX PR 24-APR-1998; 98US-0082964.
XX PA (UNIW) UNIV WASHINGTON.
XX PI Rice CM, Frolov I, McBride MS;
XX DR WPI; 2000-013359/01.
XX DR N-PSDB; AA236195.
XX PT Chimeric viral RNA, used in vaccine against BVDV -
XX PS Disclosure; Fig 10; 108pp; English.
XX CC The present sequence represents the NADL protein of bovine viral diarrhea
CC virus (BVDV), and is encoded by the low copy number plasmid pACNR/BVDV
CC NADL. The plasmid is used in the course of the invention, to produce
CC chimeric RNA viruses. The specification describes chimeric viral
CC RNA comprising a 5' nontranslated region (5'NTR); an open reading frame
CC (ORF) region; and a 3' NTR; where at least one of the regions is chimeric
CC and comprises a nucleotide sequence from a pestivirus in operable linkage
CC with a heterologous nucleotide sequence, preferably from HCV. The
CC chimeric viral RNA is replication-competent. The chimeric viral RNA
CC can be used in a method for identifying compounds having antiviral
CC activity against HCV. When the pestivirus viral nucleotide sequence is
CC from bovine viral diarrhea virus (BVDV), the chimeric viral RNA can be
CC used in a vaccine against BVDV.

XX SQ Sequence 3988 AA;
Query Match 34.7%; Score 497; DB 22; Length 3835;
Best Local Similarity 97.8%; Pred. No. 1.1e-38;
Matches 89; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 103 MCSRCQGHRRFEMDRPEKSAARYCAECNRLHPAEGDFWAESSMLGLKITYPALMDGKVY 162
DB 1384 MCSRCQGHRRFEMDRPEKSAARYCAECNRLHPAEGDFWAESSMLGLKITYPALMDGKVY 1443

QY 163 DITEWAGCQRVGISPDTHRVPHYHISFGSRIP 193
DB 1444 DITEWAGCQRVGISPDTHRVPHYHISFGSRMP 1474

Query Match 34.7%; Score 497; DB 21; Length 3988;
Best Local Similarity 97.8%; Pred. No. 1.1e-38;
Matches 89; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 103 MCSRCQGHRRFEMDRPEKSAARYCAECNRLHPAEGDFWAESSMLGLKITYPALMDGKVY 162
DB 1537 MCSRCQGHRRFEMDRPEKSAARYCAECNRLHPAEGDFWAESSMLGLKITYPALMDGKVY 1596

QY 163 DITEWAGCQRVGISPDTHRVPHYHISFGSRIP 193
DB 1597 DITEWAGCQRVGISPDTHRVPHYHISFGSRMP 1627

RESULT 12
ID AAY53616 standard; Protein; 3988 AA.
XX AC AAY53616;
XX DT 11-FEB-2000 (first entry)
XX DE Amino acid sequence of infectious BVDV NADL protein.
XX KW Pseudorevertant; RNA virus; chimera; BVDV; HCV; replication-competent;
XX KW 5' nontranslated region; 5'NTR; 3' NTR; pestivirus; antiviral;
XX KW bovine viral diarrhea virus; NADL; vaccine.
XX OS Bovine viral diarrhea virus.
XX PN WO9955366-A1.
XX PD 04-NOV-1999.
XX PF 23-APR-1999; 99WO-US08850.
XX PR 24-APR-1998; 98US-0082964.
XX PA (UNIW) UNIV WASHINGTON.
XX PI Rice CM, Frolov I, McBride MS;
XX DR WPI; 2000-013359/01.
XX DR N-PSDB; AA239596.
XX PT Chimeric viral RNA, used in vaccine against BVDV -
XX PS Disclosure; Fig 11; 108pp; English.
XX CC The present sequence represents the NADL protein of bovine viral diarrhea
CC virus (BVDV). The sequence is used in the course of the invention, to
CC produce chimeric RNA viruses. The specification describes chimeric viral
CC RNA comprising a 5' nontranslated region (5'NTR); an open reading frame
CC (ORF) region; and a 3' NTR; where at least one of the regions is chimeric
CC and comprises a nucleotide sequence from a pestivirus in operable linkage
CC with a heterologous nucleotide sequence, preferably from HCV. The
CC chimeric viral RNA is replication-competent. The chimeric viral RNA
CC can be used in a method for identifying compounds having antiviral
CC activity against HCV. When the pestivirus viral nucleotide sequence is
CC from bovine viral diarrhea virus (BVDV), the chimeric viral RNA can be
CC used in a vaccine against BVDV.

XX SQ Sequence 3988 AA;
Query Match 34.7%; Score 497; DB 21; Length 3988;
Best Local Similarity 97.8%; Pred. No. 1.1e-38;
Matches 89; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 103 MCSRCQGHRRFEMDRPEKSAARYCAECNRLHPAEGDFWAESSMLGLKITYPALMDGKVY 162
DB 1537 MCSRCQGHRRFEMDRPEKSAARYCAECNRLHPAEGDFWAESSMLGLKITYPALMDGKVY 1596

QY 163 DITEWAGCQRVGISPDTHRVPHYHISFGSRIP 193

Db 1597 DITWAGCQVRGIVSPDTHRVPCHFSGSRMP 1627

RESULT 13

ABB69643

ID ABB69643 standard; Protein; 970 AA.

XX AC ABB69643;

XX 26-MAR-2002 (first entry)

DT Drosophila melanogaster polypeptide SEQ ID NO 35721.

DE Drosophila: developmental biology; cell signalling; insecticide;

XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL13746.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

XX Disclosure; SEQ ID NO 35721; 21bp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA

CC sequences (AB101840-AB116175) and the encoded proteins

CC (AB57737-AB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 970 AA;

Query Match 29.0%; Score 416; DB 22; Length 970;

Best Local Similarity 36.6%; Pred No. 1.4e-31;

Matches 93; Conservative 45; Mismatches 80; Indels 36; Gaps 6;

QY 12 FHLVGVATASDELKAYRQLAVMVDPKXHHPRAEAFKVLRAAWDIVSNAEKRYE 71

DB 710 YSLGVFPDSSQEQIRKHYYKIAVLVHPDKXKQAGAEAFKVLQAFELIGEPENRLIYD 769

QY 72 M---KRAENELSRVNFSLKQDDLEAMNTMCMSCQKGRFEMDRPKSARYCAE 128

DB 770 QSIAETLHTEKAWTELHLLSLOLQTPMAEAANTIRCTCAQSHPRKLTFRPHYAARECAS 829

QY 129 CNRLHFAEGGFWAESSMLGKITVFALMDGKVDYDITWAGCQVRGIS---PDTHRVPH 185

DB 830 CKIRHSAKDGIWASTSMWGLRWKYLALMDGKVDYDITWANCQKALSHELSHWQVR 889

QY 186 ISFGSRIPCTGRQRAT-----PDAP-----PADLQDFLSRIFQ 219

DB 890 IVRGAQQQQQQQQQQQQQQQQQQQQHHQHPQPHDRGVHHPGGVSGVSEATLHEFLDNLYS 949

QY 220 VPPGOMPNG-NFEA 232

DB 950 ---GQHPGAHNAFA 960

RESULT 14

ABB77045

ID ABB77045 standard; Protein; 204 AA.

XX AC ABB77045;

XX 08-OCT-2002 (first entry)

XX Human protein sequence #2 from clone HNTPB82.

XX Human; HNTPB82; secreted protein; immunosuppressive; food preservative;

KW antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;

KW vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;

KW virucide; fungicide; ophthalmological; vulnery; gene therapy; ELISA;

KW radioimmunoassay; enzyme linked immunosorbent assay; autoimmune disease;

KW hyperproliferative disorder; cardiovascular disorder; angiogenesis;

KW cerebrovascular disorder; nervous system disorder; ocular disorder;

KW wound healing; food additive.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..43

FT Protein /label= Signal_peptide

FT /label= Mature_protein

XX WO200222638-A1.

XX 21-MAR-2002.

XX 17-JAN-2001; 2001WO-US01386.

XX 12-SEP-2000; 2000US-232104P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;

PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;

PI Ni J;

XX WPI; 2002-258041/30.

DR N-PSDB; ABL55114.

XX New nucleic acid molecules encoding 22 human secreted proteins for

PT diagnosing or treating e.g. autoimmune diseases, hyperproliferative

PT disorders, and cardiovascular disorders, and used as food additives or

PT preservatives -

XX Disclosure; Page 507-508; 526pp; English.

XX The sequence represents a protein sequence of the invention, encoded by

CC cDNA isolated from human clone ID HNTPB82. The invention relates to novel

CC isolated nucleic acid molecules encoding 22 human secreted proteins. The

CC proteins of the invention have immunosuppressive, antiarthritic,

CC antirheumatic, antiproliferative, cytostatic, cardiant, vasotropic,

CC cerebroprotective, nootropic, neuroprotective, antibacterial, virucide,

CC fungicide, ophthalmological, and vulnery activity. The polynucleotides

CC may have a use in gene therapy. The polynucleotides and polypeptides

CC encoded by them are used to prevent, treat or ameliorate a medical

CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,

CC chickens or sheep. The polynucleotides and polypeptides are also used in

CC diagnosing a pathological condition or susceptibility to a pathological

CC condition. The antibodies to the proteins can also be used in alleviating

CC symptoms associated with the disorders and in diagnostic immunoassays

CC e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA).

CC Disorders which are diagnosed or treated include autoimmune diseases,

CC hyperproliferative disorders, cardiovascular disorders, cerebrovascular
CC disorders, angiogenesis, nervous system disorders, infections caused by
CC bacteria, viruses and fungi and ocular disorders. The polypeptides can
CC also be used to aid wound healing and epithelial cell proliferation. The
CC polypeptides can also be used as a food additive or preservative.

XX
SQ Sequence 204 AA;
Query Match 19.7%; Score 283; DB 23; Length 204;
Best Local Similarity 79.7%; Pred. No. 1.9e-19;
Matches 55; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MAGVPEDELNPFHVLGVEATSDVELKKAYRQLAVMVHPDKNHHPRAEAFKVLRAAWDI 60
DB 130 MAGVPEDELNPFHVLGVEATSDVELKKAYRQLAVMVHPDKNHHPRAEAFKVFASSLGT 189
QY 61 VSNAEKRKE 69
DB 190 LSAMLKKRK 198

RESULT 15
ABG64926
ID ABG64926 standard; Protein; 204 AA.
XX
AC ABG64926;
XX
DT 27-AUG-2002 (first entry)
XX Human albumin fusion protein #1601.
XX
XX Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW haematopoietic disorder; neural disorder; connective disorder;
KW cytostatic; antiinfertility; antinflammatory; antiulcer;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; neurotropic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.
XX
OS Homo sapiens.
OS Synthetic.
OS
PN WO200177137-A1.
XX
XX 18-OCT-2001.
XX
XX 12-APR-2001; 2001WO-US11988.
XX
XX 12-APR-2000; 2000US-229358P.
PR 25-APR-2000; 2000US-199384P.
PR 21-DEC-2000; 2000US-256931P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Rosen CA, Haseltine WA;
PI
XX
XX WPI; 2002-010886/01.
DR
XX
XX New fusion protein for treating disease e.g. diabetes comprises an
PT albumin fused to a therapeutic protein -
XX
XX Claim 1; Page 1605-1606; 2102pp; English.
XX
XX The present invention relates to albumin fusion proteins comprising a
CC therapeutic protein X and human albumin (HA, also known as human serum
CC albumin, HSA). The proteins are useful for treating a disease or
CC disorder that may be modulated by therapeutic protein X. The albumin
CC extends the shelf-life of protein X, and may increase its biological
CC in vitro/in vivo activity. The protein is useful for treating and
CC diagnosing disorders such as cancer, reproductive disorders, digestive
CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders

CC (e.g. diabetes), haematopoietic disorders, neural disorders
CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
CC encephalomyelitis, meningitis, schizophrenia), and connective disorders
CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin
CC fusion proteins of the invention.

XX
SQ Sequence 204 AA;
Query Match 19.7%; Score 283; DB 23; Length 204;
Best Local Similarity 79.7%; Pred. No. 1.9e-19;
Matches 55; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MAGVPEDELNPFHVLGVEATSDVELKKAYRQLAVMVHPDKNHHPRAEAFKVLRAAWDI 60
DB 130 MAGVPEDELNPFHVLGVEATSDVELKKAYRQLAVMVHPDKNHHPRAEAFKVFASSLGT 189
QY 61 VSNAEKRKE 69
DB 190 LSAMLKKRK 198

Search completed: January 2, 2004, 16:40:49
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 16:38:38 ; Search time 20 Seconds

(without alignments)
1293.469 Million cell updates/sec

Title: US-10-049-742-11

Perfect score: 1433

Sequence: 1 MAGVPEELNPFHVLGVEAT.....VPKGEAKPKRKVKRRPFQR 269

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	497	34.7	3988	1 GNMVBV	genome polyprotein
2	290	20.2	815	2 T15402	hypothetical prote
3	284.5	19.9	387	2 A83302	hypothetical prote
4	282.5	19.7	751	2 T15403	hypothetical prote
5	264.5	18.5	577	2 T01052	hypothetical prote
6	183	12.8	224	2 S54519	HLJ1 protein - yea
7	169.5	11.8	409	2 S26703	dnaj protein homol
8	169	11.8	539	2 T06150	hypothetical prote
9	168.5	11.8	378	2 T24254	hypothetical prote
10	167	11.7	656	2 H84649	probable DnaJ prot
11	166.5	11.6	364	2 E70449	chaperone DnaJ - A
12	163.5	11.4	407	2 T39658	probable mitochond
13	160.5	11.2	403	2 T39697	DNAJ protein - fis
14	159	11.1	311	2 B84428	hypothetical prote
15	158	11.0	346	2 B84602	probable DnaJ prot
16	157	11.0	178	2 G02272	heat shock protein
17	157	11.0	397	2 S34630	dnaj protein homol
18	156	10.9	189	2 S34632	dnaj protein homol
19	156	10.9	384	2 A96624	hypothetical prote
20	155.5	10.9	358	2 T52073	ER-associated Hsp4
21	155	10.8	359	2 T51903	related to Hsp1 pr
22	154	10.7	374	2 S41758	heat shock protein
23	154	10.7	374	2 C97058	molecular chaperon
24	153.5	10.7	215	2 T16542	hypothetical prote
25	153.5	10.7	314	2 D87592	dnaj family protei
26	153.5	10.7	499	2 G96831	hypothetical prote
27	153	10.7	367	2 T45812	dnaj-like protein
28	153	10.7	376	2 E70361	chaperone DnaJ - A
29	152.5	10.6	427	2 B86408	probable dnaj prot

hypothetical prote
hypothetical prote
heat shock protein
DnaJ-like protein
hypothetical prote
hypothetical prote
co-chaperone-curve
heat shock protein
probable DnaJ-fam
hypothetical prote
Hsp70 cofactor [im
probable co-chaper
hypothetical prote
heat shock protein
heat-shock protein
hypothetical prote
hypothetical prote

ALIGNMENTS

RESULT 1

GNMVBV

genome polyprotein - bovine viral diarrhea virus

C:Species: bovine viral diarrhea virus, BVDV

C>Date: 30-Sep-1990 #sequence_revision 30-Sep-1990 #text_change 19-Jan-2001

C/Accession: A29198; A61161

R/Collett, M.S.; Larson, R.; Gold, C.; Strick, D.; Anderson, D.K.; Purchio, A.F.

Virolology 165, 191-199, 1988

A/Title: Molecular cloning and nucleotide sequence of the pestivirus bovine viral diarr

A/Reference number: A29198; MUID:88265858; PMID:2838957

A/Accession: A29198

A/Molecule type: genomic RNA

A/Residues: 1-3988 <COL>

A/Experimental source: isolate NADL

R/Ward, P.; Misra, V.

Am J. Vet. Res. 52, 1231-1236, 1991

A/Title: Detection of bovine viral diarrhea virus, using degenerate oligonucleotide prim

A/Reference number: A61161; MUID:92027091; PMID:1656820

A/Accession: A61161

A/Molecule type: genomic RNA

A/Residues: 2054-2072 <WAR>

A/Experimental source: isolate V1352

A/Note: authors translated the codon ATA for residue 18 as Thr

C:Superfamily: pestivirus genome polyprotein

C:Keywords: ATP; glycoprotein; nucleotide binding; P-loop; polyprotein

F:2-234/Product: viral proteinase p20 #status predicted <VPT>

F:548-1115/Product: major envelope glycoprotein gp55 #status predicted <EGP>

F:1905-1912/Region: nucleotide-binding motif A (P-loop)

F:1996-2001/Region: nucleotide-binding motif B

F:2000-2003/Region: DEXH motif

F:272,281,296,365,370,413,487,597,809,922,990,1357,1419,1451,1803,2234,2307,2584,2772,2

Query Match 34.7%; Score 497; DB 1; Length 3988;

Best Local Similarity 97.8%; Pred. No. 5.8e-30;

Matches 89; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 103 MCSFCQKHRRFFEDREPKSARYCAECNRLHPABEGDFWABSSMLGLKITYPALMDGKYY 162

Db 1537 MCSFCQKHRRFFEDREPKSARYCAECNRLHPABEGDFWABSSMLGLKITYPALMDGKYY 1596

Qy 163 DITWAGCQRVGISPDTHRVFCHISFGSRIP 193

Db 1597 DITWAGCQRVGISPDTHRVFCHISFGSRMP 1627

RESULT 2

T15402

hypothetical protein C04A2.7a - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 23-Mar-2001

C/Accession: T15402

R/Du, Z.

Db	293	DVSEWAI:COGMACRPNTHRPFSFHVNM---VGLEKATQBSKSRFPDWL	337
RESULT 4			
T15403			
hypothetical protein C04A2.7 - Caenorhabditis elegans			
C:Species: Caenorhabditis elegans			
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 23-Mar-2001			
C:Accession: T15403			
R:Du, Z.			
submitted to the EMBL Data Library, July 1995			
A:Description: The sequence of C. elegans cosmid C04A2.			
A:Reference number: S59416			
A:Accession: T15403			
A>Status: preliminary; translated from GB/EMBL/DBJ			
A:Molecule type: DNA			
A:Residues: 1-751 <DUZ>			
A:Cross-references: EMBL:U23448; NID:G733539; PID:G733543; PIDN:AAC46735.1; CESP:C04A2.7			
A:Experimental source: strain Bristol N2			
C:Genetics:			
A:Gene: CESP:C04A2.7			
A:Introns: 6/2; 73/3; 129/3; 181/2; 259/3; 382/1; 443/1; 593/3; 632/3; 688/3; 722/3; 746			
Query Match 19.7%; Score 282.5; DB 2; Length 751;			
Best Local Similarity 32.3%; Pred. No. 3.9e-14;			
Matches 52; Conservative 34; Mismatches 74; Indels 1; Gaps 1;			
QY	12	PHVLGVATASDVLEKAYROLAVMHPDKNHPRAEAFKVLRAAWDIVSNAEKREYE	71
Db	564	YSVFLGRSDCDDIKRNYKRLAALVSPDKCTTDAADQVYELVDVAFSAIGYKDSRSEYT	623
QY	72	MKMAENELSRVNEFLSKLQDDLKEAMNTMCSRCQGHRRFEMDREPKSARYCAECNR	131
Db	624	LENLKNNEVHEQLISVWMDMTKAVEEARNTIFCD-CENTHFRVATSISSQARSCKRCGV	682
QY	132	LHPAEEGDVFAESSMLGLKITFYALMDGKVDITEWAGCQ	172
Db	683	KHPAKQNDIWEKRLGLTSYYTCTDNVYDITSWATCKK	723
RESULT 5			
T01052			
hypothetical protein YUP8H12R.35 - Arabidopsis thaliana			
C:Species: Arabidopsis thaliana (mouse-ear cress)			
C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 22-Oct-1999			
C:Accession: T01052			
R:Theologis, A.; Vysotskaia, V.S.; Osborne, B.I.; Schwartz, J.R.; Federspiel, N.A.; Kwar			
Oefner, P.; Davis, B.W.			
submitted to the EMBL Data Library, May 1998			
A:Description: Arabidopsis thaliana chromosome 1 YAC YUP8H12R sequence.			
A:Reference number: Z14227			
A:Accession: T01052			
A>Status: translated from GB/EMBL/DBJ			
A:Molecule type: DNA			
A:Residues: 1-577 <THE>			
A:Cross-references: EMBL:AC002886; NID:G2494106; PID:G3152572; GSPDB:GN00059; ATSP:YUP8H			
A:Experimental source: cultivar Columbia			
C:Genetics:			
A:Gene: ATSP:YUP8H12R.35			
A:Map position: 1			
A:Introns: 107/3; 151/1; 319/3; 362/3; 397/3; 435/3; 463/3; 487/3			
Query Match 18.5%; Score 264.5; DB 2; Length 577;			
Best Local Similarity 29.0%; Pred. No. 7.2e-13;			
Matches 70; Conservative 41; Mismatches 85; Indels 45; Gaps 8;			
QY	7	DELNPFHVLGV--EATASDVLEKAYROLAVMHPDKN-HHPRAEAFKVLRAAWDIV--	61
Db	291	NSLHNYEALGVPRHKKIDAALVKKRYKXKMLVHPDKNMGSPHASEFCKLQSAVEVBL	350
QY	62	-----SNAEKREYEMKMAENELSRV--NEFLSKLQ---DOLKEAMNTM	102
Db	351	LIILNLYILKFLVLSDFVKRDRDYDQLRKEESRTRSCVQTSASHSQSGPDYRDESRI	410

A;Cross-references: GB:S74758; NID:G241522; PIDN:AAB20771.1; PID:G241523
R;Caplan, A.J.; Douglas, M.G.
J. Cell Biol. 114, 609-621, 1991
A;Title: Characterization of YDJ1: a yeast homologue of the bacterial dnaJ protein.
A;Reference number: A39659; MUID:91332099; PMID:1869583
A;Accession: A39659
A;Molecule type: DNA
A;Residues: 1-409 <CAP>
A;Cross-references: GB:X56560; NID:G4810; PIDN:CAA39910.1; PID:G4811
R;Bergez, P.; Doignon, F.; Crouzet, M.
Yeast 11, 967-974, 1995
A;Title: The sequence of a 44 420 bp fragment located on the left arm of chromosome XIV
A;Reference number: S58711; MUID:96021608; PMID:8533472
A;Accession: S58714
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-409 <BER>
A;Cross-references: EMBL:U12141; NID:G1314216; PIDN:AAA99647.1; PID:G994823
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
R;Bergez, P.; Doignon, F.; Crouzet, M.
submitted to the Protein Sequence Database, April 1996
A;Reference number: S62975
A;Accession: S62992
A;Molecule type: DNA
A;Residues: 1-409 <BEW>
A;Cross-references: EMBL:Z71340; NID:G1301940; PIDN:CAA95937.1; PID:G1301941; MIPS:YNL064c
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:YDJ1; MAS5
A;Cross-references: SGD:S0005008; MIPS:YNL064c
A;Map position: 14L
C;Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology
C;Keywords: heat shock; mitochondrion; stress-induced protein
F;6-70/Domain: dnaJ amino-terminal homology <DNJ>

Query Match 11.8%; Score 169.5; DB 2; Length 409;
Best Local Similarity 23.3%; Pred. No. 1.2e-05;
Matches 63; Conservative 35; Mismatches 71; Indels 101; Gaps 11;

QY 12 FHLVGEATASDVLELKAYRQLAVMVHPDKXHHPRAEAFKVLRAAWDIVSNAEKREYE 71
Db 8 YDILGVPTATDVEIKAYRKALKYHPDKNPSEAAEKFEASAAVEILSDPEKRDYD 67
QY 72 MKRMAENELSGD-----RSVNEFLSKLQD 93
Db 68 --QFGEDGLSGAGAGGFGGFGGDDIFSQFFGAGAGPRGPQRGKDKHEISASLE 125
QY 94 DL-----KRAMTMM-----CSRCKGKRRFEMDRE--PKSARYCAE 128
Db 126 ELYKGRATKALNKKQILCKECEGSGGKKGAVKCTSCNGQGIKF-VTRQMGPMIQRQTE 184
QY 129 CNRLHPAEEDGFWAESMLGLKITYPALMDGKVYDITWAGCQVRGIGSPDTHRVPHISF 188
Db 185 CDVCHGT--GD-----IIDPK-----DRCKSNGKKVNERKILEVHVE- 221
QY 189 GSRIPGTRGRQR-----ATPDAPPDL 210
Db 222 ----PGMKDQRIIVFKGEADQAPDVIGDV 247

RESULT 8
T06150
hypothetical protein F24J7.130 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 22-Oct-1999
R;Bevan, M.; Vitale, D.; Liguori, R.; Argiriou, A.; De Simone, V.; Bancroft, I.; Mewes,
submitted to the Protein Sequence Database, April 1999
A;Reference number: Z15493
A;Accession: T06150
A;Molecule type: DNA
A;Residues: 1-539 <BEV>
A;Cross-references: EMBL:AL021768; GSPDB:GN00062; ATSP:F24J7.130

QY 103 MCSRCQKRRFEMDREPKSARYCAECNRLHPAEGDFWAESSMLGLKITV----- 153
Db 411 HCTCKGNSHIWTCTNRTKAKRWQCQGYHQAKDGGWVE-----LKGTLFFRAHKIE 465
QY 154 ----FALMDCKVYDITWAGCQVRGIGSPDTHRVPHISFGSRIPGTRGRQRATPDAPPAD 209
Db 466 IPRAFCVCAESKIFDVSEAWICQGMACRPNTHRPSFHVNM---VGLEKTTQBSNSRFPWD 522
QY 210 L 210
Db 523 L 523

RESULT 6
S54519
HLJ1 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YM8520.10; protein YMR161w
C;Species: Saccharomyces cerevisiae
C;Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 29-Oct-1999
C;Accession: S54519; S59657; S54606
R;Hunt, S.; Bowman, S.
submitted to the EMBL Data Library, May 1995
A;Reference number: S54510
A;Accession: S54519
A;Molecule type: DNA
A;Residues: 1-224 <HUN>
A;Cross-references: GB:Z49705; EMBL:Z49700; NID:G825556; PIDN:CAA89797.1; PID:G825556; MIPS:YNL064c
A;Experimental source: strain AB972
R;Stepanek, P.; Guha, S.; Volkert, F.C.
submitted to the EMBL Data Library, January 1995
A;Description: HLJ1, a Saccharomyces cerevisiae homolog of Escherichia coli dnaJ with a
A;Reference number: S59657
A;Accession: S59657
A;Molecule type: DNA
A;Residues: 1-224 <STE>
A;Cross-references: EMBL:U19358; NID:G972935; PIDN:AAA75025.1; PID:G972936
C;Genetics:
A;Gene: SGD:HLJ1
A;Cross-references: SGD:S0004771; MIPS:YMR161w
A;Map position: 13R
C;Superfamily: dnaJ amino-terminal homology
C;Keywords: transmembrane protein
F;21-85/Domain: dnaJ amino-terminal homology <DNJ>
F;202-218/Domain: transmembrane #status predicted <TM>

Query Match 12.8%; Score 183; DB 2; Length 224;
Best Local Similarity 50.0%; Pred. No. 5.2e-07;
Matches 33; Conservative 17; Mismatches 16; Indels 0; Gaps 0;

QY 6 EDELNPFHVLGEATASDVLELKAYRQLAVMVHPDKXHHPRAEAFKVLRAAWDIVSNAE 65
Db 17 KDKHEFYELKVDKATDSEIKAYRKALKLHPDKXNSHPKAGEAPKVINRAFEVLSNEE 76
QY 66 KRKEYE 71
Db 77 KRKIYD 82

RESULT 7
S26703
dnaJ protein homolog YDJ1 - yeast (Saccharomyces cerevisiae)
N;Alternate names: dnaJ protein homolog MAS5; protein N2418; protein YNL064c; protein YN
C;Species: Saccharomyces cerevisiae
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Jun-2000
C;Accession: S26703; A39659; S58714; S62992; S17250
R;Atencio, D.P.; Yaffe, M.P.
Mol. Cell. Biol. 12, 283-291, 1992
A;Title: MAS5, a yeast homolog of dnaJ involved in mitochondrial protein import.
A;Reference number: S26703; MUID:92107179; PMID:1729605
A;Accession: S26703
A;Molecule type: DNA
A;Residues: 1-409 <ATE>

A;Experimental source: cultivar Columbia; BAC clone F24J7
C;Genetics:
A;Gene: ATSP:F24J7.130
A;Map position: 4
A;Introns: 242/2

Query Match 11.8%; Score 169; DB 2; Length 539;
Best Local Similarity 42.1%; Pred. No. 1.8e-05;
Matches 32; Conservative 17; Mismatches 27; Indels 0; Gaps 0;
QY 8 ELNPFHVLGVEATASDELKAYROLAVMVHPDKNHHPRAEAFKVLRAAWDIVSNAEKR 67
DB 64 EADWYRLGVDPDLADDEAVKRYKRLALLHPDKNFTGAECAFKLILEANDLLSDKSQR 123
QY 68 KEYEMKMAENELSR 83
DB 124 SSYQKRSNQVKQRT 139

RESULT 9
T24254
hypothetical protein R74.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 08-Sep-2000
C;Accession: T24254
R;Gardner, A.
submitted to the EMBL Data Library, August 1994
A;Reference number: Z19864
A;Accession: T24254
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-378 <WIL>
A;Cross-references: EMBL:Z36238; PIDN:CAA85274.1; GSPDB:GN00021; CESP:R74.4
A;Experimental source: clone R74
C;Genetics:
A;Gene: CESP:R74.4
A;Map position: 3
A;Introns: 27/3; 102/1; 157/3; 332/3
C;Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology
F;17-82/Domain: dnaJ amino-terminal homology <DNJ>

Query Match 11.8%; Score 168.5; DB 2; Length 378;
Best Local Similarity 23.3%; Pred. No. 1.3e-05;
Matches 62; Conservative 43; Mismatches 92; Indels 69; Gaps 9;
QY 2 AGVPED-----ELNPFHVLGVEATASDELKAYROLAVMVHPDKN-HHPRAEAFKVL 55
DB 4 ATTPGQPKVSEMDFYQLLGVEMKMAENELSRKAYRKALKYHDPNDNAHAEFEKVS 63
QY 56 AAWDIVSNAEKREYEMKMAENELSR-----RSVNEFLSKLQDDLKEAMNTMCSR 106
DB 64 IAYSVLDPNKRQYDVGSPSENLQDFEGFDVSEMGGVGRVFGALFSLKGVPIPTQIVPK 123
QY 107 CGKHRRFEMDREPKSARYC-AECNRLHPAE-----EGDFW----- 141
DB 124 VLAQARHICWQE-----CDVQARQLPGETVTSSVSKOHAHYEYINIOEHRKNGVAI 177
QY 142 -ASSMLGLKITFALMDGVYDITEWAGCQRVIGISPTHRVPHISFGSRIPGTRQR 200
DB 178 ICKSSSKFLVLFD-KEGGVMTQESGKRGAGTQADMFFVPYV----- 222
QY 201 ATPDAPADLQDELSTRIFQVPPGOMP 226
DB 223 -----ANTQEFNAMKYLHEDKETP 241

RESULT 10
H84649
probable DnaJ protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: H84649
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84649
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-656 <STO>
A;Cross-references: GB:AE002093; NID:96598507; PIDN:AAF18620.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g25560
A;Map position: 2

Query Match 11.7%; Score 167; DB 2; Length 656;
Best Local Similarity 23.4%; Pred. No. 3.2e-05;
Matches 68; Conservative 41; Mismatches 114; Indels 68; Gaps 9;
QY 8 ELNPFHVLGVEATASDELKAYROLAVMVHPDKNHHPRAEAFKVLRAAWDIVSNAEKR 67
DB 64 DVDHYGVGLNPEADDEIVRKRYKRLAVMLHPDRNKSVGAEAEAFKLSQAWGVFSDKAKR 123
QY 68 KEYEMK-----MAENELSRVNEFLSKLQDDLKEAMNTMCSRCCQKHRRFEMDREPKSAR 124
DB 124 ADYDLKRNVLGYKGGGASSRSPATNGFQVTKASGNTTKYKSKRGIKR-----ASDASAA 179
QY 125 YCAECNRLHPAEEDGDEWAESSMLGLKITYPALMDGVYDITEWAGCQRVIGISPTH----- 180
DB 180 ATTSTAQKTTADGTFTVTCRTCTQTEYHSVY---LNQNLCPNCRKPFIAVETDPGS 236
QY 181 ---RVPHY-ISFGSRIPGTRGRORATPDAPPADLQDFLSRIFQVPPQGMNGFN----- 230
DB 237 GSIRKTFHEHFQDLSLHTTDRKKV-----PGRDNNGVYGEYDSF 277
QY 231 ----FAAPQAPAGAAAASK-----PNSTVPKGEAKPKRRK 261
DB 278 EWGVFTGKTSAHATPTGSRKDEVRREYTKRTAGPSSTIP-----PKRRK 323

RESULT 11
E70449
Chaperone DnaJ - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 20-Aug-1999
C;Accession: E70449
R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov, V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: E70449
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-364 <AQF>
A;Cross-references: GB:AE000753; NID:92984035; PIDN:AAC07578.1; PID:92984036; GB:AE00065
A;Experimental source: strain VFS
C;Genetics:
A;Gene: dnaJ1
C;Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology
F;7-71/Domain: dnaJ amino-terminal homology <DNJ>

Query Match 11.6%; Score 166.5; DB 2; Length 364;
Best Local Similarity 24.2%; Pred. No. 1.7e-05;
Matches 62; Conservative 40; Mismatches 97; Indels 57; Gaps 10;
QY 12 FHLVGVATASDELKAYROLAVMVHPDKNHHPRAEAFKVLRAAWDIVSNAEKRKEYE 71
DB 9 YEILGVNRDATKEEIKKAYRKLVIYHPDIPDPDPAQEKFKINEAHYHLIDDERSEYD 68
QY 72 MKRVAENELSRVNEFLSKLQDDLKEAMNTMCSRCCQKHRRFEMDREPKSARYCAECNR 131
DB 69 -ALLSRNDVNGK-FRDFLEYIEFVESITQGE-----KGKRRPRKQDIDKMLPLT---- 117

QY 132 LHPAEGDFWAESSMLGLKITYFALMDGKVDITWAGCQVRGIGSPDTHRPVPHISFGSR 191
DB 118 ---LEAAGLGCCK---EIIISRWMDCPV-----CEGMVKGEAETVCHACNG-- 159
QY 192 IFGTGRORATPDAPPADLQDLFLSRIFQVP-PQOMPNGNFFFAAPQAPAGAAAASKPNSTV 250
DB 160 ---EGRR-----VSGIFNRPSPCVCKGKGFVKNPCP-----TCY 192
QY 251 PKGEAKPKRKVKRRP 266
DB 193 GRGRVSAQHKIKVHIP 208

RESULT 12
T39658
probable mitochondrial protein import protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 02-Sep-2000
C:Accession: T39658
R:Lyne, M.; Rajandream, M.A.; Bartrell, B.G.; Rieger, M.
submitted to the EMBL Data Library, October 1998
A:Reference number: Z21868
A:Accession: T39658
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-407 <LYN>
A:Cross-references: EMBL:AL031856; PIDN:CAA21305.1; GSPDB:GN000067; SPDB:SPBCL1734.11
C:Genetics:
A:Gene: dnaJ
C:Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology
C:Keywords: mitochondrion
F:6-68/Domain: dnaJ amino-terminal homology <DNJ>

Query Match 11.4%; Score 163.5; DB 2; Length 407;
Best Local Similarity 28.0%; Pred. No. 3.4e-05;
Matches 46; Conservative 23; Mismatches 44; Indels 51; Gaps 6;

QY 8 ELNPFHVLGVETASDELKAYRQLAVMVHPDKNHHPRABEAFKVLRAAMDIVNAEKR 67
DB 4 ETKLYEVLNVDTASQALKKAYRKLALKYHPDKN--PNAGDKFKIISRAVEILADESKR 61
QY 68 KEYEMKMAENELS-----RSVNEFLSKLQ---D 93
DB 62 ATYD-RFGEGLQGAGDGGMSADDLFAEFFGGMGPRKGDVLVHTTKLTLE 119
QY 94 DLKAMNT-----WMCSCQGHRRFEMDREPKSARYCAECN 130
DB 120 DLYRGKTKLALOKKVICPKCSGR-----GGKESVSKASCN 157

RESULT 13
T39697
DNAJ protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Sep-2000
C:Accession: T39697
R:Wood, V.; Skelton, J.; Churcher, C.M.; Rajandream, M.A.; Bartrell, B.G.
submitted to the EMBL Data Library, July 1999
A:Reference number: Z21870
A:Accession: T39697
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-403 <WOO>
A:Cross-references: EMBL:AL09652; PIDN:CAB51764.1; GSPDB:GN000067
A:Experimental source: strain 972h; cosmid c17A3
C:Genetics:
A:Gene: pi042
A:Map position: 2
A:Introns: 10/2; 156/1
C:Superfamily: dnaJ amino-terminal homology
F:113-177/Domain: dnaJ amino-terminal homology <DNJ>

Query Match 11.2%; Score 160.5; DB 2; Length 403;

Best Local Similarity 24.0%; Pred. No. 5.7e-05;
Matches 54; Conservative 39; Mismatches 63; Indels 69; Gaps 10;

QY 12 FHVLGVEATASDELKAYRQLAVMVHPDKNHHPRABEAFKVLRAAMDIVNAEKRYE 71
DB 115 YEILDLLKTKTCTDEIKSKYKLLALQLHPDKOHAPSADAEAFKWSKAFQVLSDPNLAHYD 174
QY 72 MKRM-AENELSRSVNEFLSKL-----QDDLKEAM----- 99
DB 175 RTGMDPESRASAASSSFSSNAGGHPFSAYPQANMSPEDLNFSGDQOFFSGPGTFFGG 234
QY 100 -NTWMCSCQGHRRF-----EMDREPKSARYCAECNLRHP-----AEEGDF-WAESS 145
DB 235 GPGIRVHQFGGRPRNFARROQAQDMPPKSIFY-----QLLPLIVVILFAFLSNFSWSDST 289
QY 146 MLGLKITYFALMDGKVDITWAGCQVRGIGSPDT---HRVPYHIS 187
DB 290 SVN---TRYSFQQNYKTV-----PRTTAKHNIPYMS 319

RESULT 14
B84428
hypothetical protein At2g01710 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: B84428
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-769, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: B84428
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-311 <STO>
A:Cross-references: GB:AE002093; NID:G4220477; PIDN:AAD12700.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g01710
A:Map position: 2

Query Match 11.1%; Score 159; DB 2; Length 311;
Best Local Similarity 21.9%; Pred. No. 5.6e-05;
Matches 68; Conservative 46; Mismatches 85; Indels 112; Gaps 16;

QY 1 MAGVPEDEL---NPFHVLGVE-----ATASDELKAYRQLAVMVHPDKNHHPRABEAF 51
DB 58 LSSAPENRIKQNPWYKILQIEDLTESSTDNDL-IKKQYRRLALLLHPDKNRPFPADQAF 116
QY 52 KVLRAAMDIVNAEKRYEYEMKMAENELSR-----SVNEFLSKLQDDLKEAMNTM--MC 104
DB 117 RFVLDAWEVLSTPTKKSQF-----DGLNLIFTKVNLNTQSKKKKTTTNEKMTFTWTAC 170
QY 105 SRCQGHRRFEMDREPKSARYCAECNLRHPABEGDFWAESSMLGLKITYFALMDGKYDI 164
DB 171 PYCYSHEVPRVYQE-----YCIRCO----- 191
QY 165 TEWAGCQVRGIGSPDTHRP-----YHISFG---SRIFGTGRORATPDAPPADLQDF 213
DB 192 ----NCQRAFHAASIQPLPLIPGKDEYCCWGGFFPMGFGVGGKGEAAI--ANGVDAAKF 245
QY 214 LSRIFQVPPQOMP---NGNFFAAPQAPGAA-----AASKPNSTV----- 250
DB 246 -----PNWMPFVSSGGVAAPPSSGNGVSGVFDGWSGGAARKRDNEAVRSNNGVGVNSDGT 297
QY 251 PK--GEAKPKR 259
DB 298 PKRGRGRPKK 308

RESULT 15
B84602
probable DnaJ protein [imported] - Arabidopsis thaliana

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 16:37:16 ; Search time 17 Seconds
(without alignments)
744.128 Million cell updates/sec

Title: US-10-049-742-11

Perfect score: 1433
Sequence: 1 MAGVPEELNPFHVLGVEAT.....VPKGEAKPKRKKVRRPQR 269

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	497	34.7	3988	1 POLG_BVDVN	P19711 bovine vira
2	290	20.2	915	1 YQ07_CABEL	Q09446 caenorhabdi
3	193	13.5	376	1 DJBC_MOUSE	Q9qy14 mus musculu
4	189	13.2	375	1 DJBC_HUMAN	Q9nkw2 homo sapien
5	183	12.8	224	1 HLJ1_YEAST	P48353 saccharomyc
6	172.5	12.0	412	1 DJJ2_HUMAN	O60884 homo sapien
7	172.5	12.0	412	1 DJA2_MOUSE	Q9qy10 mus musculu
8	171.5	12.0	412	1 DJA2_RAT	O35824 rattus norv
9	169.5	11.8	409	1 MARS_YEAST	P25491 saccharomyc
10	166.5	11.6	364	1 DJJ1_AQUAE	O67623 aquifex aeo
11	163.5	11.4	387	1 DNAJ_METTE	Q9uxr9 methanosarc
12	162	11.3	397	1 DJR4_HUMAN	Q8wv22 homo sapien
13	160.5	11.2	337	1 DJB4_MOUSE	Q9d832 mus musculu
14	160.5	11.2	370	1 DNAJ_EYRHR	O05646 erysipelothe
15	159	11.1	334	1 DNJ1_DROME	Q24133 drosophila
16	157	11.0	397	1 DJA1_HUMAN	P31689 homo sapien
17	157	11.0	397	1 DJA4_MOUSE	Q9jmc3 mus musculu
18	156.5	10.9	280	1 DNAJ_TRETH	Q56237 thermus the
19	156.5	10.9	337	1 DJB4_HUMAN	Q9udy4 homo sapien
20	156.5	10.9	348	1 DJB5_MOUSE	O89114 mus musculu
21	155.5	10.9	358	1 DJBB_HUMAN	Q9ub54 homo sapien
22	155.5	10.9	358	1 DJBB_MOUSE	Q99kvl mus musculu
23	155.5	10.9	371	1 DNAJ_METSS	Q9zfc5 methylovoru
24	154	10.7	348	1 DJB5_HUMAN	O75953 homo sapien
25	154	10.7	374	1 DNAJ_CLOAB	P30725 clostridium
26	153	10.7	375	1 DNAJ_BRECH	Q91c44 brevivacill
27	153	10.7	376	1 DNJ2_AQUAE	O66921 aquifex aeo
28	151	10.5	397	1 DJA1_MOUSE	P54102 mus musculu
29	150	10.5	379	1 DNJ2_STRAL	O52184 streptomyc
30	149.5	10.4	223	1 DJB9_HUMAN	Q9ub53 homo sapien
31	149.5	10.4	383	1 DJB9_PORGI	Q9xca6 porphyromon
32	149.5	10.4	552	1 DJC1_MOUSE	O61712 mus musculu
33	149	10.4	222	1 DJB9_RAT	P97554 rattus norv

34	149	10.4	389	1 DNAJ_METMA	P35515 methanosarc
35	148.5	10.4	590	1 YNW7_YEAST	P53863 saccharomyc
36	148	10.3	378	1 DNJ2_STRCO	Q9rd7 streptomyc
37	148	10.3	388	1 DNAJ_MYCLE	Q02605 mycobacteri
38	147.5	10.3	332	1 DNJH_SYNV3	P50027 synchocyst
39	147	10.3	232	1 DJB8_HUMAN	Q8nho0 homo sapien
40	147	10.3	340	1 DJB1_HUMAN	P25885 homo sapien
41	147	10.3	554	1 DJC1_HUMAN	Q96kc8 homo sapien
42	146.5	10.2	376	1 DNJ8_METTH	O27352 methanobact
43	146	10.2	264	1 DJC8_HUMAN	O75937 homo sapien
44	145.5	10.2	480	1 DJA3_HUMAN	Q96ev1 homo sapien
45	145	10.1	370	1 DNAJ_BACHD	Q9kd71 bacillus ha

ALIGNMENTS

RESULT 1

POLG_BVDVN	STANDARD;	PRT; 3988 AA.
AC	P19711;	
DT	01-FEB-1991 (Rel. 17, Created)	
DT	01-FEB-1996 (Rel. 33, Last sequence update)	
DT	01-FEB-1996 (Rel. 33, Last annotation update)	
DE	Genome polyprotein.	
OS	Bovine viral diarrhea virus (isolate NADL) (BVDV) (Mucosal disease virus).	
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Pestivirus.	
OX	NCBI_TaxID=111100;	
RN	[1]_TaxID=111100;	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=88265859; PubMed=2838957;	
RA	Collett M.S., Larson R., Gold C., Strick D., Anderson D.K., Purchio A.F.;	
RT	"Molecular cloning and nucleotide sequence of the pestivirus bovine viral diarrhea virus";	
RL	Virology 165:191-199(1988).	
RN	[2]	
RP	GENOMIC ORGANIZATION.	
RX	MEDLINE=88265859; PubMed=2838958;	
RA	Collett M.S., Larson R., Belzer S.K., Retzel E.;	
RT	"Proteins encoded by bovine viral diarrhea virus: the genomic organization of a pestivirus.";	
RL	Virology 165:200-208(1988).	
CC	-1- FUNCTION: PESTIVIRUS P80 (P125) MAY BE A BIFUNCTIONAL PROTEIN	
CC	-1- WITH HELICASE AND PROTEASE ACTIVITY.	
CC	-1- PTM: GP116 GIVES RISE TO GP62 AND GP53; GP62 IN TURN YIELDS GP48 AND GP25.	
CC	-1- SIMILARITY: TO THE HOG CHOLERA VIRUS GENOME POLYPROTEIN.	
CC	-1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S31.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
CC	-----	
EMBL; M31182; AAA42854.1; ..		
PIR; A29198; GNWVBV.		
DR	HSP; P27958; 1A1V.	
DR	MEROPS; C53.001; ..	
DR	MEROPS; C51.001; ..	
DR	InterPro; IPR000280; CDvir_endptsep80.	
DR	InterPro; IPR001410; DEAD.	
DR	InterPro; IPR002166; HCV RdRP.	
DR	InterPro; IPR001650; Helicase_C.	
DR	InterPro; IPR007095; RNA_pol_DS_PS.	
DR	InterPro; IPR007094; RNA_pol_PSVir.	
DR	InterPro; IPR001568; RNase_T2.	
DR	Pfam; PF00271; helicase_C; 1.	

```
DR PFam: PF00998; Viral RdRP; 1.
DR PRINTS: PR00729; CDVENDOPTASE.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00490; HELIC; 1.
DR PROSITE: PS00531; RNASE T2.2; UNKNOWN 1.
KW Polypeptide; Glycoprotein; Helicase; Serine protease; Hydrolase.
FT CHAIN 1 7270
FT CHAIN 2 71063
FT CHAIN 3 3988
FT CHAIN 4 272
FT CARBOHYD 272
FT CARBOHYD 281
FT CARBOHYD 296
FT CARBOHYD 335
FT CARBOHYD 365
FT CARBOHYD 370
FT CARBOHYD 413
FT CARBOHYD 487
FT CARBOHYD 597
FT CARBOHYD 809
FT CARBOHYD 878
FT CARBOHYD 922
FT CARBOHYD 990
FT CARBOHYD 1357
FT CARBOHYD 1419
FT CARBOHYD 1451
FT CARBOHYD 1803
FT CARBOHYD 2224
FT CARBOHYD 2307
FT CARBOHYD 2584
FT CARBOHYD 2772
FT CARBOHYD 2981
FT CARBOHYD 3778
FT CARBOHYD 3867
FT CARBOHYD 3883
SQ SEQUENCE 3988 AA; 449154 MW; 4474212F338661B8 CRC64;

Query Match 34.7%; Score 497; DB 1; Length 3988;
Best Local Similarity 97.8%; Pred. No. 5.2e-31;
Matches 89; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 103 MCSRCQGGKRRFEMDRPKSARYCAECNRLHPAEEGDFWAESSMLGKITYFALMDGKVY 162
DB 1537 MCSRCQGGKRRFEMDRPKSARYCAECNRLHPAEEGDFWAESSMLGKITYFALMDGKVY 1596

QY 163 DITEWAGCQGVISPDTHRVPHYHISFGSRIP 193
DB 1597 DITEWAGCQGVISPDTHRVPHYHISFGSRMP 1627

RESULT 2
YQ07 CAEEL STANDARD; PRT; 915 AA.
AC Q09416; Q95QY2; Q95QY3; Q95QY4;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C04A2.7 in chromosome II.
GN C04A2.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Du Z.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS, AND ALTERNATIVE SPLICING.
RA Waterston R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -!- ALTERNATIVE PRODUCTS;
```

```
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Experimental confirmation may be lacking for some
CC isoforms;
CC Name=A;
CC IsoId=Q09446-1; Sequence=Displayed;
CC Name=B;
CC IsoId=Q09446-2; Sequence=VSP_001299, VSP_001301;
CC Name=C;
CC IsoId=Q09446-3; Sequence=VSP_001298, VSP_001300;
CC -!- SIMILARITY: Contains 1 J domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; U23448; AAL27225.1; -
CC EMBL; U23448; AAL27226.1; -
CC EMBL; U23448; AAL27227.1; -
CC WormPep; C04A2.7a; CE29657.
CC WormPep; C04A2.7b; CE29658.
CC WormPep; C04A2.7c; CE29659.
CC InterPro; IPR001623; DnaJ_N.
CC Pfam; PF00226; DnaJ; 1.
CC SMART; SM00271; DnaJ; 1.
CC PROSITE; PS00636; DnaJ_1; FALSE_NEG.
CC PROSITE; PS00076; DnaJ_2; 1.
CC Hypothetical protein; Alternative splicing.
CC DOMAIN 660 724 J-DOMAIN.
CC DOMAIN 65 70 POLY-PRO.
CC DOMAIN 137 164 HIS-RICH.
CC VARSPLIC 160 160 Missing (in isoform C).
CC VARSPLIC 821 849 Missing (in isoform B).
CC VARSPLIC 845 846 Missing (in isoform C).
CC VARSPLIC 850 915 Missing (in isoform B).
CC SEQUENCE 915 AA; 102739 MW; 1B38BD9DC4133C1 CRC64;

Query Match 20.2%; Score 290; DB 1; Length 915;
Best Local Similarity 32.2%; Pred. No. 2.3e-15;
Matches 58; Conservative 34; Mismatches 82; Indels 6; Gaps 3;

QY 12 PHVLGVETASDVVELKAYROLAVNHPDQNHPRAEAFKVLRAAWDIVSNAEKKEYE 71
DB 662 YSVFGLRSDGSDDDIKRNYKRLAALVSPDKCTIDAADQVVELVDVAFSAIGYKDSRSEY 721

QY 72 MKRMAENELSRVNEFLSKLQDLKEAMNTVMGSCQCKHRRFEMDRPKSARYCAECNR 131
DB 722 LENLKNNEVEHQLISVNMNDMTKAVEEARNTIFCD-CENTHFRVATISPSQARSKRCGV 780

QY 132 LHPAEEGDFWAESSMLGKITYFALMDGKVYDITEWAGC--QRV---GISPDTHRVPHYH 186
DB 781 KHPAKQNDIWEKRLHGLTSTYYCTDNNVYDITSWATCKSQRAMLKNMRAHTNVPYRL 840

RESULT 3
DUBC_MOUSE STANDARD; PRT; 376 AA.
AC Q9QYI4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DnaJ homolog subfamily B member 12 (mDj10).
GN DnaJB12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```



```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR BMBL; AK000034; BAA90896.1; .
DR HSP; P25685; 1HDJ
DR Genew; HGNC:14891; DNAJB12.
DR InterPro; IPR0011623; DnaJ_N.
DR InterPro; IPR003095; Hsp_DnaJ.
DR Pfam; PF00226; DnaJ; 1.
DR PRINTS; PRO0625; DNAJPROTEIN.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS00636; DNAJ_1; 1.
DR PROSITE; PS0076; DNAJ_2; 1.
DR Chaperone; Transmembrane.
KW Chaperone; Transmembrane.
FT DOMAIN 112 176 J-DOMAIN.
FT TRANSMEM 244 264 POTENTIAL.
SQ SEQUENCE 375 AA; 41785 MW; EBF393EF44AE67B4 CRC64;
Query Match 13.2%; Score 189; DB 1; Length 375;
Best Local Similarity 38.3%; Pred. No. 7.9e-08;
Matches 44; Conservative 14; Mismatches 27; Indels 30; Gaps 3;
QY 12 FHLVGVGVEATPSDELKAYRQLAVMVHPDKNHHPRABEAFKVLRRAWDIVLSNAEKRYE 71
DB 112 YELTGVSGASDEDLKAYRRLLAKLFHDPKNHAPGATEAFKAIGTAYAVLSNPKRKQYD 171
QY 72 MKRMAENELSRSVNFELSKLODDLEAMNTMCSRCSQGK-----HRRPEMDREPX 121
DB 172 -----QFGDDKSQAAR-----HGHHGHGDHFHRGFADISPE 201
RESULT 5
HLJI_YEAST STANDARD; PRF; 224 AA.
AC P48353;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE HLJ1 protein.
GN HLJ1 OR YMR161W OR YMB520.10.
OS Saccharomyces cerevisiae (Baker's yeast)..
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OC NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XIII."
RL Nature 387:90-93(1997).
CC -I- SIMILARITY: Contains 1 J domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; U19358; AAA75025.1; -
DR EMBL; Z49705; CAA89737.1; -
DR PIR; S54519; S54519.
DR HSSP; P25685; 1HDJ.
DR SGD; S0004771; HLJ1.
DR InterPro; IPR001623; DnaJ_N.
DR InterPro; IPR003095; Hsp_DnaJ.
DR Pfam; PF00226; DnaJ; 1.
DR PRINTS; PR00625; DNAJPROTEIN.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS00636; DnaJ_1; 1.
DR PROSITE; PS50076; DnaJ_2; 1.
KW Chaperone.
FT DOMAIN 18 87 J-DOMAIN.
SQ SEQUENCE 224 AA; 25008 MW; A9BFED9D242C2DD CRC64;

Query Match 12.8%; Score 183; DB 1; Length 224;
Best Local Similarity 50.0%; Pred. No. 1.3e-07;
Matches 33; Conservative 17; Mismatches 16; Indels 0; Gaps 0;

QY 6 EDELNPFHVLGVEATSDVELKKAYQLAVMHPDKNHPRAEEAFKVLRAAWDIVSNAE 65
DB 17 KKHFEYELKVDKATDSEIKKAYKLAIKLHPDKNSHPKAGEAFKVINRAPEVLSNEE 76

QY 66 KRKEYE 71
DB 77 KRSIYD 82

RESULT 6
ID DJA2 HUMAN STANDARD; PRT; 412 AA.
AC O60984; O14711.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE DnaJ homolog subfamily A member 2 (HIRA interacting protein 4) (Cell
DE cycle progression restoration gene 3 protein) (DnaJ3).
GN DNaJ2 OR HIRIP4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RA Lorain S., Brendel C., Scamps C., Lecluse Y., Lipinski M.;
RT "HIRIP4, a new human DnaJ, is a nuclear protein that interacts with
RT the product of the DiGeorge syndrome gene candidate HIRA.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98043401; PubMed=9383053;
RA Edwards M.C., Liegeois N., Horecka J., Depinho R.A., Sprague G.F. Jr.,
RA Tyers M., Elledge S.J.;
RT "Human CPR (cell cycle progression restoration) genes impart a Far-
RT phenotype on yeast cells";
RL Genetics 147:1063-1076(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta, and Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
```

```
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smaalus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=20387370; PubMed=10816573;
RA Terada K., Mori M.;
RT "Human DnaJ homologs dj2 and dj3, and bag-1 are positive cochaperones
RT of hsc70.";
RL J. Biol. Chem. 275:24728-24734(2000).
CC -!- FUNCTION: Co-chaperone of Hsc70.
CC -!- SUBUNIT: INTERACTS WITH HIRA.
CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND (POTENTIAL) OR NUCLEAR.
CC -!- PM: Farnesylated.
CC -!- SIMILARITY: Contains 1 J domain.
CC -!- SIMILARITY: Contains 1 CR domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ001309; CAA04669.1; -
DR EMBL; Y13350; CAA73791.1; -
DR EMBL; AF011793; AAB69313.1; -
DR EMBL; BC013044; AAH13044.1; -
DR EMBL; BC015809; AAH15809.1; -
DR HSSP; P25685; 1HDJ.
DR Genew; HGNC:14884; DNAJ2.
DR GO; GO:0008284; P:positive regulation of cell proliferation; TAS.
DR GO; GO:0000074; P:regulation of cell cycle; TAS.
DR InterPro; IPR002939; DnaJ_C.
DR InterPro; IPR001305; DnaJ_CXXCXGKG.
DR InterPro; IPR001623; DnaJ_N.
DR InterPro; IPR003095; Hsp_DnaJ.
DR Pfam; PF00226; DnaJ; 1.
DR Pfam; PF01556; DnaJ_C; 1.
DR Pfam; PF00684; DnaJ_CXXCXGKG; 1.
DR PRINTS; PR00625; DNAJPROTEIN.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS00636; DnaJ_1; 1.
DR PROSITE; PS50076; DnaJ_2; 1.
DR PROSITE; PS00637; DnaJ_CXXCXGKG; 1.
KW Chaperone; Repeat; Zinc; Metal-binding; Prenylation; Lipoprotein;
KW Membrane; Multigene family.
FT DOMAIN 8 70 J-DOMAIN.
FT REPEAT 143 150 CXXCXGKG MOTIF.
FT REPEAT 159 166 CXXCXGKG MOTIF.
FT REPEAT 186 193 CXXCXGKG MOTIF.
FT REPEAT 202 209 CXXCXGKG MOTIF.
FT METAL 143 143 ZINC 1 (BY SIMILARITY).
FT METAL 146 146 ZINC 1 (BY SIMILARITY).
FT METAL 159 159 ZINC 2 (BY SIMILARITY).
FT METAL 162 162 ZINC 2 (BY SIMILARITY).
FT METAL 186 186 ZINC 2 (BY SIMILARITY).
FT METAL 189 189 ZINC 2 (BY SIMILARITY).
FT METAL 202 202 ZINC 1 (BY SIMILARITY).
FT METAL 205 205 ZINC 1 (BY SIMILARITY).
FT LIPID 409 409 FARNESYL (BY SIMILARITY).
FT CONFLICT 17 17 P -> A (IN REF. 2).
FT CONFLICT 42 46 NAGDK -> QNQETN (IN REF. 2).
FT CONFLICT 83 93 GMDIFSHIFG -> WHGLIFSLTVFC (IN REF. 2).
FT CONFLICT 242 257 GVPEGDIVLLQKEKH -> EWNPTLFFLLPGKNM (IN REF. 2).
```

```

FT CONFLICT 286 287 FK -> LS (IN REF. 2).
FT CONFLICT 328 328 D -> G (IN REF. 2).
SQ SEQUENCE 412 AA; 45745 MW; 8F1BC367425CB428 CRC64;

Query Match
Best Local Similarity 12.0%; Score 172.5; DB 1; Length 412;
Matches 83; Conservative 44; Mismatches 89; Indels 163; Gaps 17;

Qy 1 MAGVPEDELNPHVLGVTEASDVLEKAYRQLAVMVHPDKNHHPRAEAFKVLRAAWDI 60
Db 1 MANVADTKL--YDILGVPPGASENELKAYRKLAKKEYHPDKN--PNAGDKFKEISFAYEV 56
Qy 61 VSNAEKREKEYEMKMAENEL-----SRSVN----- 85
Db 57 LSNPEKRELYD--RYGEOQLREGSGGGGMDIDFISHIFGGGLFGFMGNQSRNRGRERGE 114
Qy 86 EFLSKQLDDLKEAMN-----TMCSCRCQGHRRFEMDREPKSAR----- 124
Db 115 DMHPLKVSLELDYNGKTKTLQLSKNVLCSCSGQGGKSGAVQKCSACRGVRIMIRQL 174
Qy 125 -----YCAECNRLHPAEEGDFWABSSML-----GLKITVFALMDGKVYDITEWAGC 170
Db 175 APMVQMQQSCVSCDCN-----GEGEVINEKDRCKCEKKV-----IKEVKILEV----- 219
Qy 171 QRVGISPDTHRVPHYHISFGSRIPGTRGRATPDAPPAD-----LQDFLSRIFQ----- 219
Db 220 -----HVDKGMKHGQRTFT-GEADQAPGVEPGDVLVLLQEKHEVFRDGNDLH 268
Qy 220 -----VPPGQMPNGNFFAPOPAPGAGAAASXPNSTV 250
Db 269 MTKIGLVEALCGFQTFKHLDRQIVVKYPPGVIE-----PGCV-----RV 311
Qy 251 PKGEAKPKRKKVRPPFOR 269
Db 312 VRGEGMP-----QYRNPFEK 326

RESULT 7
DUA2 MOUSE STANDARD; PRT; 412 AA.
AC Q9QYJ0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE DnaJ homolog subfamily A member 2 (mdj3).
GN DUAJ2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=21023480; PubMed=11147971;
RA Ohtsuka K., Hata M.;
RT "Mammalian HSP40/DNAJ homologs: cloning of novel cDNAs and a proposal
RT for their classification and nomenclature.";
RL Cell Stress Chapterones 5:98-112(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=2338257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

```

```

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences.";
CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Co-chaperone of Hsc70 (By similarity).
CC -!- SUBCELLULAR LOCATION: Membrane-bound (Potential).
CC -!- PTM: Farnesylated (By similarity).
CC -!- SIMILARITY: Contains 1 J domain.
CC -!- SIMILARITY: Contains 1 CR domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB028853; BAA88301.1; -.
DR EMBL; BC003420; AAH03420.1; -.
DR HSSP; P25685; 1HDJ.
DR MGD; MGI:1931882; DnaJ2.
DR InterPro; IPR002919; DnaJ_C.
DR InterPro; IPR001105; DnaJ_CXXCXGXG.
DR InterPro; IPR001623; DnaJ_N.
DR InterPro; IPR003095; Hsp_DnaJ.
DR Pfam; PF00226; DnaJ; 1.
DR Pfam; PF01556; DnaJ_C; 1.
DR Pfam; PF00684; DnaJ_CXXCXGXG; 1.
DR PRINTS; PRO0625; DnaJPROTEIN.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS00636; DnaJ_1; 1.
DR PROSITE; PS00076; DnaJ_2; 1.
DR PROSITE; PS00637; DnaJ_CXXCXGXG; 1.
KW Chaperone; Repeat; Zinc; Metal-binding; Prenylation; Lipoprotein;
KW Membrane; Multigene family.
FT DOMAIN 8 70 J-DOMAIN.
FT REPEAT 143 150 CXXCXGXG MOTIF.
FT REPEAT 159 166 CXXCXGXG MOTIF.
FT REPEAT 186 193 CXXCXGXG MOTIF.
FT REPEAT 202 209 CXXCXGXG MOTIF.
FT METAL 143 143 ZINC 1 (BY SIMILARITY).
FT METAL 146 146 ZINC 1 (BY SIMILARITY).
FT METAL 159 159 ZINC 2 (BY SIMILARITY).
FT METAL 162 162 ZINC 2 (BY SIMILARITY).
FT METAL 186 186 ZINC 2 (BY SIMILARITY).
FT METAL 189 189 ZINC 2 (BY SIMILARITY).
FT METAL 202 202 ZINC 1 (BY SIMILARITY).
FT METAL 205 205 ZINC 1 (BY SIMILARITY).
FT LIPID 409 409 FARNESYL (BY SIMILARITY).
SQ SEQUENCE 412 AA; 45745 MW; 98130EC0925CB42E CRC64;

Query Match
Best Local Similarity 12.0%; Score 172.5; DB 1; Length 412;
Matches 83; Conservative 44; Mismatches 89; Indels 163; Gaps 17;

Qy 1 MAGVPEDELNPHVLGVTEASDVLEKAYRQLAVMVHPDKNHHPRAEAFKVLRAAWDI 60
Db 1 MANVADTKL--YDILGVPPGASENELKAYRKLAKKEYHPDKN--PNAGDKFKEISFAYEV 56
Qy 61 VSNAEKREKEYEMKMAENEL-----SRSVN----- 85
Db 57 LSNPEKRELYD--RYGEOQLREGSGGGGMDIDFISHIFGGGLFGFMGNQSRNRGRERGE 114
Qy 86 EFLSKQLDDLKEAMN-----TMCSCRCQGHRRFEMDREPKSAR----- 124
Db 115 DMHPLKVSLELDYNGKTKTLQLSKNVLCSCSGQGGKSGAVQKCSACRGVRIMIRQL 174
Qy 125 -----YCAECNRLHPAEEGDFWABSSML-----GLKITVFALMDGKVYDITEWAGC 170

```

Db 175 APMVQVQMSVSCDCN-----GEGEVINEKDRCKCKGKVV-----IKEVKILEV----- 219
QY 171 QRVGISPDPTRVPHVHISFGSRIPGTRGRQATPDAPPAD-----LQDFLSRIFQ----- 219
Db 220 -----HVDKGMKHGQRIITFT-GEADQAPGVEPGDIVLLLOEKEHEVFQDGDNDLH 268
QY 220 -----VPPQMPNGNFFAAPQAPGAAAAKPNSTV 250
Db 269 MTYKIGLVEALCGFQFTFKHLDARQIVVKYPPGKVIE-----PGCV-----RV 311
QY 251 PKGEAKPKRKKVRRPQOR 269
Db 312 VRGEGMP-----QYRNPFEK 326

RESULT 8
ID DJA2 RAT STANDARD; PRT; 412 AA.
AC O35824;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE DnaJ homolog subfamily A member 2 (RDJ2).
GN DNAJ2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97466951; PubMed=9328291;
RA Andres D.A., Shao H., Crick D.C., Finlin B.S.;
RT "Expressional cloning of a novel farnesylated protein, RDJ2, encoding a
DnaJ protein homologue."
RL Arch. Biochem. Biophys. 346:113-124(1997).
CC -!- SURCELLULAR LOCATION: Membrane-bound (Potential).
CC -!- SIMILARITY: Contains 1 J domain.
CC -!- SIMILARITY: Contains 1 CR domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U95727; AAB64094.1; -
DR HSP; P25685; 1HDJ.
DR InterPro; IPR002939; DnaJ_C.
DR InterPro; IPR001305; DnaJ_CXXCXGXG.
DR InterPro; IPR001623; DnaJ_N.
DR InterPro; IPR003095; Hsp_DnaJ.
DR Pfam; PF00226; DnaJ; 1.
DR Pfam; PF01556; DnaJ_C; 1.
DR Pfam; PF00684; DnaJ_CXXCXGXG; 1.
DR PRINTS; PR00625; DNAJPROTEIN.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS00636; DnaJ_1; 1.
DR PROSITE; PS0076; DnaJ_2; 1.
DR PROSITE; PS00637; DnaJ_CXXCXGXG; 1.
KW Chapterone; Repeat; Zinc; Metal-binding; Prenylation; Lipoprotein;
KW Multigene family.
KW DOMAIN 8
FT REPEAT 143 150 J-DOMAIN.
FT REPEAT 159 166 CXXCXGXG MOTIF.
FT REPEAT 186 193 CXXCXGXG MOTIF.
FT REPEAT 202 209 CXXCXGXG MOTIF.
FT METAL 143 143 ZINC 1 (BY SIMILARITY).
FT METAL 146 146 ZINC 1 (BY SIMILARITY).
FT METAL 159 159 ZINC 2 (BY SIMILARITY).
FT METAL 162 162 ZINC 2 (BY SIMILARITY).

FT METAL 186 186 ZINC 2 (BY SIMILARITY).
FT METAL 189 189 ZINC 2 (BY SIMILARITY).
FT METAL 202 202 ZINC 1 (BY SIMILARITY).
FT METAL 205 205 ZINC 1 (BY SIMILARITY).
FT LIPID 409 409 FARNESYL (BY SIMILARITY).
SQ SEQUENCE 412 AA; 45765 MW; FAE60601B55C409F CRC64;
Query Match 12.0%; Score 171.5; DB 1; Length 412;
Best Local Similarity 21.6%; Pred. No. 2.2e-06;
Matches 82; Conservative 45; Mismatches 89; Indels 163; Gaps 17;
QY 1 MAGVPEDELNPFLVGLVEATASDVLEKAYRQCAVMVHPDKNHHPRAEAEAFKVLRAAWDI 60
Db 1 MANVADTKL--YDILGVPPGASENELKAYRKLAKYHPDKN--PNAGDKFEISFAYEV 56
QY 61 VSNAEKKEKYEKMAENEL-----SRSVN----- 85
Db 57 LSNPEKRELYD--RYGEQGLREGSGGGGMDIDFHSIFGGGLFGFMGNQSRNRGRRE 114
QY 86 EFLSKLQDDIKEAMN-----TWCSCRCQCKHRRFEMDRPKSAR----- 124
Db 115 DMHPLKVSLELYNGTKTKLQSKNVLCSCSGGKSGAVQKCSACRGRGVRIMIRQL 174
QY 125 -----YCAECNRLHPAEEDFWAESML-----GLKITYFALMDGKVVYDITEWAGC 170
Db 175 APMVQVQMSVSCDCN-----GEGEVINEKDRCKCKGKVV-----IKEVKILEV----- 219
QY 171 QRVGISPDPTRVPHVHISFGSRIPGTRGRQATPDAPPADLQDFLS-----RIFQ----- 219
Db 220 -----HVDKGMKHGQRIITFT-GEADQAPGVEPGDIVLVQKEHEVFQDGDNDLH 268
QY 220 -----VPPQMPNGNFFAAPQAPGAAAAKPNSTV 250
Db 269 MTYKIGLVEALCGFQFTFKHLDARQIVVKYPPGKVIE-----PGCV-----RV 311
QY 251 PKGEAKPKRKKVRRPQOR 269
Db 312 VRGEGMP-----QYRNPFEK 326
RESULT 9
ID MASS_YEAST STANDARD; PRT; 409 AA.
AC P25491;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitochondrial protein import protein MAS5 (Protein YDJ1).
GN MAS5 OR YDJ1 OR YNL084C OR N2418 OR YNL2418C.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92107179; PubMed=1729605;
RA Atencio D.P., Yaffe M.P.;
RT "MAS5, a yeast homolog of DnaJ involved in mitochondrial protein
import."
RL Mol. Cell. Biol. 12:283-291(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91332099; PubMed=1869583;
RA Caplan A.J., Douglas M.G.;
RT "Characterization of YDJ1: a yeast homologue of the bacterial dnaJ
protein."
RL J. Cell Biol. 114:609-621(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / FY1676;
RX MEDLINE=96021608; PubMed=8533472;
RA Bergez P., Doignon F., Crouzet M.;
RT "The sequence of a 44 420 bp fragment located on the left arm of

FT REPEAT 137 144 CXXCXGXG MOTIF.
FT REPEAT 154 161 CXXCXGXG MOTIF.
FT REPEAT 174 181 CXXCXGXG MOTIF.
FT REPEAT 188 195 CXXCXGXG MOTIF.
FT METAL 137 137 ZINC 1 (BY SIMILARITY).
FT METAL 140 140 ZINC 1 (BY SIMILARITY).
FT METAL 154 154 ZINC 2 (BY SIMILARITY).
FT METAL 157 157 ZINC 2 (BY SIMILARITY).
FT METAL 174 174 ZINC 2 (BY SIMILARITY).
FT METAL 177 177 ZINC 1 (BY SIMILARITY).
FT METAL 188 188 ZINC 1 (BY SIMILARITY).
FT METAL 191 191 ZINC 1 (BY SIMILARITY).
SQ SEQUENCE 364 AA; 40937 MW; ABD709EE59B0E7E7 CRC64;

Query Match 11.6%; Score 166.5; DB 1; Length 364;
Best Local Similarity 24.2%; Pred. No. 4.7e-06;
Matches 62; Conservative 40; Mismatches 97; Indels 57; Gaps 10;

QY 12 FHVLGVEATSDVELKAYROLAVMHPDKXHHPRAEAEAFKVLRAAWDIVSNAEKRYE 71
DB 9 YEILGVRDATBEIKKAYRKLRIYHPDINPPSAQEKREINAYHVILIDDERSEYD 68

QY 72 MKGMAENELSRVNEFLSKLQDOLKEAMNTMCSRQCGKRRFEMDRPKSARYCAECNR 131
DB 69 -AILSRNDVGK-PRDFLEYIQEFVESIIQGE-----KGKKRPRKGGDIKWLPLT---- 117

QY 132 LHPAEGDFWAESSMLGLKITFYFALMDGKYYDITWAGQGVGSPDTHRVPHISFGSR 191
DB 118 ---LEEAGLCEK-----ELIYSRWDCPV-----CEGMVGKEATVVCACNG-- 159

QY 192 IPGTRGRQATPDAPPADLOFLSRIFQVP-PCGMPNGNFFAAPPQAPGAAAKSPNSTV 250
DB 160 ----EGR-----VSGIFNFRPCSVCKGKGFVKNPCP-----TCY 192

QY 251 PKGEAKPKRRKKVRRP 266
DB 193 GRGRVSAQHKIKVHP 208

RESULT 11
DNBJ METTE STANDARD; PRT; 387 AA.
ID DNBJ METTE STANDARD; PRT; 387 AA.
AC Q9UXR9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chaperone protein dnaJ (Heat shock protein 40).
GN DNBJ OR HSP40.
OS Methanosarcina thermophila.
OC Archaea, Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1825 / TM-1;
RX MEDLINE=20035743; PubMed=10570966;
RA Hofman-Bang J.P., Lange M., Conway de Macario E., Macario A.J.P.,
RA Ahning B.K.;
RT "The genes coding for the hsp70 (dnaK) molecular chaperone machine
RT occur in the moderate thermophilic archaeon Methanosarcina thermophila
RT TM-1";
RL Gene 238:387-395(1999).
CC -!- FUNCTION: ACTS AS A CO-CHAPERONE. STIMULATES, JOINTLY WITH GRPE,
CC -!- THE ATPASE ACTIVITY OF DNAK (BY SIMILARITY).
CC -!- COFACTOR: BINDS TWO ZINC IONS PER MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE DNABJ FAMILY.
CC -!- SIMILARITY: Contains 1 J domain.
CC -!- SIMILARITY: Contains 1 CR domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AJ010152; CAB53763.1; -.
DR HSP; P25685; 1HDJ.
DR InterPro; IPR002939; DnaJ_C.
DR InterPro; IPR001305; DnaJ_CXXCXGXG.
DR InterPro; IPR001623; DnaJ_N.
DR InterPro; IPR003095; Hsp_DnaJ.
DR Pfam; PF00226; DnaJ; 1.
DR Pfam; PF01556; DnaJ_C; 1.
DR Pfam; PF00884; DnaJ_CXXCXGXG; 1.
DR PRINTS; PRO0625; DNAJPROTEIN.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS00636; DnaJ_1; 1.
DR PROSITE; PS00076; DnaJ_2; 1.
DR PROSITE; PS00637; DnaJ_CXXCXGXG; 1.
KW Chaperone; DNA replication; Heat shock; Repeat; Zinc; Metal-binding.
FT DOMAIN 6 70 J-DOMAIN.
FT DOMAIN 73 109 GLY-RICH.
FT REPEAT 143 150 CXXCXGXG MOTIF.
FT REPEAT 160 167 CXXCXGXG MOTIF.
FT REPEAT 186 193 CXXCXGXG MOTIF.
FT REPEAT 200 207 CXXCXGXG MOTIF.
FT METAL 143 143 ZINC 1 (BY SIMILARITY).
FT METAL 146 146 ZINC 1 (BY SIMILARITY).
FT METAL 160 160 ZINC 2 (BY SIMILARITY).
FT METAL 163 163 ZINC 2 (BY SIMILARITY).
FT METAL 186 186 ZINC 2 (BY SIMILARITY).
FT METAL 189 189 ZINC 2 (BY SIMILARITY).
FT METAL 200 200 ZINC 1 (BY SIMILARITY).
FT METAL 203 203 ZINC 1 (BY SIMILARITY).
SQ SEQUENCE 387 AA; 42554 MW; 4FDA50214424904D CRC64;

Query Match 11.4%; Score 163.5; DB 1; Length 387;
Best Local Similarity 25.0%; Pred. No. 8.7e-06;
Matches 63; Conservative 34; Mismatches 90; Indels 65; Gaps 10;

QY 12 FHVLGVEATSDVELKAYROLAVMHPDKXHHPRAEAEAFKVLRAAWDIVSNAEKRYE 70
DB 8 YEILGVRDATPEIDIKSYRKLALYHPDRNKEPCAEKFEISEAYAVLSDPKRAQYD 67

QY 71 -----EMKMAENELSRVNEFL-----SKLQDDL-----KEA 98
DB 68 RFGHAGINGQYTAEDIFRGAFGSGDIFEMFGSGRRGPRGPRGSDLYDLYITFEA 127
QY 99 M-----NTMCSRQCGKRRFEMDRPKSARYCAECNRLHPAEGDFWAESSMLG 148
DB 128 AFGVRKIDIVPRTERCSCNCSGTGAR--PGTSPKRCPTCGTGQIRTRTG-----LG 177
QY 149 LK1---TYFALMDGKYYDI-TWAGQGVGISPDPDTHRVPHISFGS-----RIPGTRGR 198
DB 178 MQFVSTTTCSTCRGKGQVIESPCVCSGTGRVTRTKITVNPAGADSGMSLRLSGE--G 235

RESULT 12
DNBJ_HUMAN
ID DNBJ_HUMAN STANDARD; PRT; 397 AA.
AC Q8W22; Q8N7P2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE DnaJ homolog subfamily A member 4.
GN DNABJ4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC TISSUE=Brain, and Trachea;
 RA Nishi T., Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H.,
 RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
 RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
 RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
 RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
 RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
 RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
 RA "NEDO human cDNA sequencing project";
 RL Submitted (JUL-2002) to the ENBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=22386257; PubMed=12477932;
 RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RL 1- SUBCELLULAR LOCATION: Membrane-bound (Potential).
 CC 1- SIMILARITY: Contains 1 J domain.
 CC 1- SIMILARITY: Contains 1 CR domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR ENBL; AK096616; BAC04828.1;
 DR ENBL; AK098079; BAC05229.1; ALT_INIT.
 DR ENBL; BC021720; AAH21720.1; -;
 DR Genew; HGNC:14885; DNaJ.A4.
 DR InterPro; IPR002939; DnaJ_C.
 DR InterPro; IPR001305; DnaJ_CXXCXXGXG.
 DR InterPro; IPR001623; DnaJ_N.
 DR InterPro; IPR003095; Hsp_DnaJ.
 DR Pfam; PF00226; DnaJ; 1.
 DR Pfam; PF00684; DnaJ_CXXCXXGXG; 1.
 DR Pfam; PF01556; DnaJ_C; 1.
 DR PRINTS; PR00825; DNaJPROTEIN.
 DR PROSITE; PS00636; DnaJ_1; 1.
 DR PROSITE; PS00737; DnaJ_2; 1.
 DR PROSITE; PS00637; DnaJ_CXXCXXGXG; 1.
 KW Chaperone; Repeat; Zinc; Metal-binding; Prenylation; Lipoprotein;
 KW Multigene family.
 FT DOMAIN 4 70 J-DOMAIN.
 FT DOMAIN 75 96 GLY-RICH
 FT REPEAT 135 142 CXXCXXGXG MOTIF.
 FT REPEAT 151 158 CXXCXXGXG MOTIF.
 FT REPEAT 178 185 CXXCXXGXG MOTIF.
 FT REPEAT 194 201 CXXCXXGXG MOTIF.
 FT REPEAT 201 201 CXXCXXGXG MOTIF.
 FT METAL 135 ZINC 1 (BY SIMILARITY).

FT METAL 138 ZINC 1 (BY SIMILARITY).
 FT METAL 151 ZINC 2 (BY SIMILARITY).
 FT METAL 154 ZINC 2 (BY SIMILARITY).
 FT METAL 178 ZINC 2 (BY SIMILARITY).
 FT METAL 181 ZINC 2 (BY SIMILARITY).
 FT METAL 194 ZINC 1 (BY SIMILARITY).
 FT METAL 197 ZINC 1 (BY SIMILARITY).
 FT LIPID 394 FARNESYL (BY SIMILARITY).
 FT CONFLICT 33 Y -> C (IN REF. 1; BAC05229).
 SQ SEQUENCE 397 AA; 44798 MW; 60D6AAD12C9B2529 CRC64;
 Query Match 11.3%; Score 162; DB 1; Length 397;
 Best Local Similarity 21.5%; Pred No. 1.2e-05;
 Matches 57; Conservative 44; Mismatches 74; Indels 90; Gaps 10;
 QY 8 ELNPHVLGVGATASDELKAYRQLAVMVHPDKNHHPRABEAFKVLRAWDIVSNAER 67
 DB 4 ETQYDILGVKPSAPSEIKKAYRKLKLYHPDKN--PDEGEKFLKLSQAYEVLSPKXR 61
 QY 68 KEYEM-----KRAENELSRVNEFLSKLQDLKE 97
 DB 62 DYVDQGGQAIKEGGSGSPSPMDIFDMFFGGGRWARERRGKNVHQLSVTLDELIN 121
 QY 98 AM-----NTMCSRQCG-----KRRPFEMDRE-----PKSARYCAEC 129
 DB 122 GVTKLALQKNVICKEGCVGKGSVEKCPCKGRGMQIHQIQIGMVGQIQTVCIET 181
 QY 130 ----NRLHFAEEDGFWESEMLGLKITYPALMDGKYVDITETWAGCORVGISPTHVRPYH 185
 DB 182 KGGERINPDRCCE-----SCGAKV-----IREKKIIEVHVKEGGMK-----DQKILFH 226
 QY 186 ISFGSRIPGTRGRQRATPPADL 210
 DB 227 -----GEGDQEPLEPGDV 240
 RESULT 13
 DJB4 MOUSE
 ID DJB4 MOUSE STANDARD; PRT; 337 AA.
 AC Q9D832; Q9D9U2;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DnaJ homolog subfamily B member 4.
 GN DNaJB4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 OX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; Tissue=Small intestine, and Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Mateu Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).

DNJ1 DROME
ID DNJ1 DROME STANDARD; PRT: 334 AA.
AC Q24133; Q9VRP0;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE DnaJ protein homolog 1 (DROJ1).
GN DnaJ-1 OR DROJ1 OR CG10578.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee J.Y., Palter K.B.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkeley; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.P., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Botkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fodor C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkeley; TISSUE=Embryo;
RA MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brockstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource";
RL Genome Biol. 3:RESEARCH080.1-RESEARCH080.8(2002).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- INDUCTION: By heat shock.
CC -!- SIMILARITY: Contains 1 J domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; U34904; AAC23584.1; -;
DR EMBL; AE003565; AAF50753.1; -;
DR EMBL; AY058788; AAL14017.1; -;
DR HSSP; P25685; 1HDJ.
DR FlyBase; FBgn0015657; DnaJ-1.
DR InterPro; IPR002939; DnaJ_C.
DR InterPro; IPR001823; DnaJ_N.
DR InterPro; IPR003095; Hsp_DnaJ.
DR Pfam; PF00226; DnaJ; 1.
DR Pfam; PF01556; DnaJ_C; 1.
DR PRINTS; PR00625; DnaJPROTEIN.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS00636; DnaJ_1; 1.
DR PROSITE; PS0076; DnaJ_2; 1.
KW Chapterone; Heat shock.
FT DOMAIN 261 261 G -> E (IN REF. 1).
FT CONFLICT 4 68 J-DOMAIN.
SQ SEQUENCE 334 AA; 37028 MW; 60542ABFD47A5689 CRC64;
Query Match 11.1%; Score 159; DB 1; Length 334;
Best Local Similarity 48.3%; Pred. No. 1.6e-05;
Matches 29; Conservative 16; Mismatches 15; Indels 0; Gaps 0;
QY 12 FHVLGVEATASDELKAYRQLAVMTVPDKNHHPRAEBAFKVLRAAWDIVSNAEKRYE 71
Db 6 YKILGLERKASDDDEIKKAYRKLALKYHPDKNKSPOAEERFKETAEYEVLSDKKRDIFD 65
Search completed: January 2, 2004, 16:41:19
Job time : 18 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 16:40:23 ; Search time 25 Seconds

(without alignments)
455.265 Million cell updates/sec

Title: US-10-049-742-11

Perfect score: 1433

Sequence: 1 MAGVPEDELNPFHVLGVEAT.....VPGKAPKPKKVRPFQR 269

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	173.5	12.1	375	4	US-09-328-352-4984
2	157	11.0	397	2	US-08-868-288A-5
3	157	11.0	397	3	US-09-235-373-5
4	157	11.0	397	3	US-09-388-993-5
5	155.5	10.9	358	2	US-08-868-288A-1
6	155.5	10.9	358	3	US-09-235-373-1
7	155.5	10.9	358	3	US-09-388-993-1
8	155.5	10.9	358	4	US-09-996-243-148
9	154	10.7	348	2	US-08-974-546-1
10	149.5	10.4	223	4	US-09-658-644-4
11	147	10.3	340	2	US-08-974-546-5
12	147	10.3	419	2	US-08-686-417-3
13	146	10.2	265	4	US-09-149-476-497
14	144	10.0	127	4	US-09-370-838-199
15	143.5	10.0	381	4	US-09-252-991A-27174
16	143.5	10.0	493	4	US-09-613-303-19
17	141.5	9.9	392	4	US-09-198-452A-48
18	140	9.8	131	4	US-09-553-498-4
19	140	9.8	131	4	US-09-618-869-4
20	140	9.8	399	4	US-09-553-498-2
21	140	9.8	399	4	US-09-618-869-2
22	137	9.6	320	4	US-09-328-352-7937
23	134	9.4	51	1	US-08-346-849-12
24	134	9.4	51	2	US-08-293-284A-12
25	134	9.4	51	4	US-08-898-300-12
26	134	9.4	330	2	US-08-868-288A-3
27	134	9.4	330	3	US-09-235-373-3

28	134	9.4	330	3	US-09-388-993-3	Sequence 3, Appli
29	133	9.3	385	4	US-09-134-001C-3688	Sequence 3688, Ap
30	126.5	8.8	352	2	US-08-472-534-6	Sequence 6, Appli
31	125.5	8.8	87	4	US-09-882-835-4	Sequence 4, Appli
32	124.5	8.7	332	4	US-09-882-835-2	Sequence 2, Appli
33	119.5	8.3	172	4	US-09-293-549-12	Sequence 12, Appli
34	119.5	8.3	351	2	US-08-868-288A-6	Sequence 6, Appli
35	119.5	8.3	351	3	US-09-235-373-6	Sequence 6, Appli
36	119.5	8.3	351	3	US-09-388-993-6	Sequence 6, Appli
37	118	8.2	438	2	US-08-897-340-34	Sequence 34, Appli
38	118	8.2	438	3	US-09-252-329-34	Sequence 34, Appli
39	118	8.2	484	2	US-08-879-260-4	Sequence 4, Appli
40	118	8.2	484	3	US-09-231-529-4	Sequence 4, Appli
41	118	8.2	484	3	US-08-977-816-4	Sequence 4, Appli
42	114	8.0	277	2	US-08-868-288A-7	Sequence 7, Appli
43	114	8.0	277	3	US-09-235-373-7	Sequence 7, Appli
44	114	8.0	277	3	US-09-388-993-7	Sequence 7, Appli
45	108.5	7.6	52	1	US-08-346-849-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-328-352-4984
; Sequence 4984, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4984
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (322)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknc
US-09-328-352-4984

Query Match	12.1%	Score 173.5;	DB 4;	Length 375;
Best Local Similarity	24.1%	Pred. No. 2.5e-09;		
Matches	63;	Conservative 31;	Mismatches 84;	Indels 83; Gaps 11;
Qy	12	FHVLGVEATASDVELKAYRQLAVMHPDKN-HHPRAEEAFKVLRAAMDIVSNAEKREY	70	
Db	12	YEVLGVSXTASDDEIKKAYRKLAMKYHDPNPDNAEAEKFKAESEAYEILSDSEKRSY	71	
Qy	71	EMKRNVAENLSR-----SVNEFLSKLD-----DLKEAM	99	
Db	72	D--RMGNAFEGGCGAGGCGFSAEDIFSGIFGAGFGGGRQQRGRGSLRYM	129	
Qy	100	NTWM-----CSHCQKHRRFFENDREPKSARYCAECNRLHPAEGDFW	141	
Db	130	ELTLEEAAYGVKKTTFTAPAPCDYCDGK-----SKPKDVETCKTCH-----	173	
Qy	142	RESSMLGKITVYFALM-----DGKYYDITEMAGCORVGISPDTH-----RVPYHISFG	189	
Db	174	-GSGGVRRVQGGFFSVQQTGCTGCGGKLIK-NPCHACHGSGVADRQQTLEVTIPAGVDNG	231	
Qy	190	SRIPGTRGRQRPATPDAPPADL	210	
Db	232	DRV-RLSGKGEAIRDQAGDL	251	

RESULT 2
US-08-868-288A-5
; Sequence 5, Application US/08868288A

```

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/235,373
  FILING DATE:
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 08/868,288
    FILING DATE: June 3, 1997
  ATTORNEY/AGENT INFORMATION:
    NAME: Billings, Lucy J.
    REGISTRATION NUMBER: 36,749
    REFERENCE/DOCKET NUMBER: PF-0309 US
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: 415-855-0555
      TELEFAX: 415-845-4166
    INFORMATION FOR SEQ ID NO: 5:
      SEQUENCE CHARACTERISTICS:
        LENGTH: 397 amino acids
        TYPE: amino acid
        STRANDEDNESS: single
        TOPOLOGY: linear
      IMMEDIATE SOURCE:
        LIBRARY: GenBank
        CLONE: 306714
  US-09-235-373-5

Query Match          11.0%; Score 157; DB 3; Length 397;
Best Local Similarity 26.0%; Pred. No. 1.3e-07;
Matches 45; Conservative 25; Mismatches 53; Indels 50; Gaps 5;

QY      8 ELNPHVLGVGVEATSDVELKAYROLAVMVHPDKNHPRAEEAFKYLRAAWDIVSNAEKR 67
DB      4 ETTYDVLGVKPNATQEEELKKAYRLALKYHPDKN--PNEGEKFKQISQAYEVLSDAKKR 61

QY      68 KEYEM-----KMAENELSRSVNEFLSKLQDOLKEA 98
DB      62 ELYDKGGEQAIKEGGAGGGFGSPMDIFDMFFGGGGRMRERRGNVYVHQLSVTLDELNG 121

QY      99 -----MNTMMSRCQGKHRRFEMDRPKSARYCAECN-----RLHPAEEG 138
DB      122 ATRKLALQKNVICDKCEGR-----GGKKGAVECCPCRCGTGMQIRIHQIGPG 168

RESULT 4
US-09-388-993-5
; Sequence 5, Application US/09388993
; Patent No. 6043222
; GENERAL INFORMATION:
;   APPLICANT: Au-Young, Janice
;   APPLICANT: Lal, Preeti
;   APPLICANT: Bandman, Olga
;   TITLE OF INVENTION: TWO NEW HUMAN DNAJ-LIKE PROTEINS
;   NUMBER OF SEQUENCES: 7
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Incyte Pharmaceuticals, Inc.
;     STREET: 3174 Porter Drive
;     CITY: Palo Alto
;     STATE: CA
;     COUNTRY: USA
;     ZIP: 94304
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette
;     COMPUTER: IBM Compatible
;     OPERATING SYSTEM: DOS
;     SOFTWARE: FastSeq for Windows Version 2.0
;     CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/09/388,993
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/868,288
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0309 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 397 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 306714
;
US-09-388-993-5
;
Query Match 11.0%; Score 157; DB 3; Length 397;
Best Local Similarity 26.0%; Pred. No. 1.3e-07;
Matches 45; Conservative 25; Mismatches 53; Indels 50; Gaps 5;

QY 8 ELNPHVLGVATASDVELKAYROLAVMVHPDKN-HHPRAEEAFKVLRAAMDIVSNAEKR 67
DB 4 ETTYDVLGVPNATQBELKAYRKALQYHPDKN--PNEGKFKQISQAYEVLSDAKKR 61
QY 68 KEYEM-----KRMENELSRVNEFLSKLODDLKEA 98
DB 62 ELYDKGGCAIKEGAGGGGSPMIDFMFGGGGRMQRERRGKNVHQLSVTLEDLYNG 121
QY 99 -----MNTMWCSCQGHRRFEMDRPKSARYCAECN-----RLHPAEG 138
DB 122 ATRKLALQGVICDRCEGR-----GGKGAVECCPNCRCGTGMQIRHQIGFG 168

RESULT 5
US-08-868-288A-1
; Sequence 1, Application US/08868288A
; Patent No. 5922567
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Lal, Preeti
; APPLICANT: Bandman, Olga
; TITLE OF INVENTION: TWO NEW HUMAN DNAJ-LIKE PROTEINS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: June 3, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0309 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 358 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYNCRAB01
; CLONE: 136466
;
US-08-868-288A-1
;
Query Match 10.9%; Score 155.5; DB 3; Length 358;
Best Local Similarity 47.5%; Pred. No. 1.6e-07;
Matches 29; Conservative 20; Mismatches 11; Indels 1; Gaps 1;

QY 12 FHVLGVEATASDVELKAYROLAVMVHPDKN-HHPRAEEAFKVLRAAMDIVSNAEKR 70
DB 27 YKILGVPRASIKDKKAYRKALQYHPDRPDPPQAEKFDLGAAYEVLSDSEKRQY 86
QY 71 E 71
DB 87 D 87

RESULT 6
US-09-235-373-1
; Sequence 1, Application US/09235373
; Patent No. 6001598
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Lal, Preeti
; APPLICANT: Bandman, Olga
; TITLE OF INVENTION: TWO NEW HUMAN DNAJ-LIKE PROTEINS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/235,373
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/868,288
; FILING DATE: June 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0309 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 358 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYNCRAB01
; CLONE: 136466
;
US-09-235-373-1
;
Query Match 10.9%; Score 155.5; DB 3; Length 358;
Best Local Similarity 47.5%; Pred. No. 1.6e-07;
Matches 29; Conservative 20; Mismatches 11; Indels 1; Gaps 1;

QY 12 FHVLGVEATASDVELKAYROLAVMVHPDKN-HHPRAEEAFKVLRAAMDIVSNAEKR 70
DB 27 YKILGVPRASIKDKKAYRKALQYHPDRPDPPQAEKFDLGAAYEVLSDSEKRQY 86
QY 71 E 71
DB 87 D 87
```


1	PRIOR FILING DATE: 1998-06-04	
2	PRIOR APPLICATION NUMBER: 607/088326	
3	PRIOR FILING DATE: 1998-06-04	
4	PRIOR APPLICATION NUMBER: 607/088167	
5	PRIOR FILING DATE: 1998-06-05	
6	PRIOR APPLICATION NUMBER: 607/088202	
7	PRIOR FILING DATE: 1998-06-05	
8	PRIOR APPLICATION NUMBER: 607/088212	
9	PRIOR FILING DATE: 1998-06-05	
10	PRIOR APPLICATION NUMBER: 607/088217	
11	PRIOR FILING DATE: 1998-06-05	
12	PRIOR APPLICATION NUMBER: 607/088655	
13	PRIOR FILING DATE: 1998-06-09	
14	PRIOR APPLICATION NUMBER: 607/088734	
15	PRIOR FILING DATE: 1998-06-10	
16	PRIOR APPLICATION NUMBER: 607/088738	
17	PRIOR FILING DATE: 1998-06-10	
18	PRIOR APPLICATION NUMBER: 607/088742	
19	PRIOR FILING DATE: 1998-06-10	
20	PRIOR APPLICATION NUMBER: 607/088910	
21	PRIOR FILING DATE: 1998-06-10	
22	PRIOR APPLICATION NUMBER: 607/088824	
23	PRIOR FILING DATE: 1998-06-10	
24	PRIOR APPLICATION NUMBER: 607/088826	
25	PRIOR FILING DATE: 1998-06-10	
26	PRIOR APPLICATION NUMBER: 607/088858	
27	PRIOR FILING DATE: 1998-06-11	
28	PRIOR APPLICATION NUMBER: 607/088861	
29	PRIOR FILING DATE: 1998-06-11	
30	PRIOR APPLICATION NUMBER: 607/088876	
31	PRIOR FILING DATE: 1998-06-11	
32	PRIOR APPLICATION NUMBER: 607/089105	
33	PRIOR FILING DATE: 1998-06-12	
34	PRIOR APPLICATION NUMBER: 607/089440	
35	PRIOR FILING DATE: 1998-06-16	
36	PRIOR APPLICATION NUMBER: 607/089512	
37	PRIOR FILING DATE: 1998-06-16	
38	PRIOR APPLICATION NUMBER: 607/089514	
39	PRIOR FILING DATE: 1998-06-16	
40	PRIOR APPLICATION NUMBER: 607/089532	
41	PRIOR FILING DATE: 1998-06-17	
42	PRIOR APPLICATION NUMBER: 607/089538	
43	PRIOR FILING DATE: 1998-06-17	
44	PRIOR APPLICATION NUMBER: 607/089598	
45	PRIOR FILING DATE: 1998-06-17	
46	PRIOR APPLICATION NUMBER: 607/089599	
47	PRIOR FILING DATE: 1998-06-17	
48	PRIOR APPLICATION NUMBER: 607/089600	
49	PRIOR FILING DATE: 1998-06-17	
50	PRIOR APPLICATION NUMBER: 607/089653	
51	PRIOR FILING DATE: 1998-06-17	
52	PRIOR APPLICATION NUMBER: 607/089801	
53	PRIOR FILING DATE: 1998-06-18	
54	PRIOR APPLICATION NUMBER: 607/089907	
55	PRIOR FILING DATE: 1998-06-18	
56	PRIOR APPLICATION NUMBER: 607/089908	
57	PRIOR FILING DATE: 1998-06-18	
58	PRIOR APPLICATION NUMBER: 607/089952	
59	PRIOR FILING DATE: 1998-06-19	
60	PRIOR APPLICATION NUMBER: 607/090254	
61	PRIOR FILING DATE: 1998-06-22	
62	PRIOR APPLICATION NUMBER: 607/090246	
63	PRIOR FILING DATE: 1998-06-22	
64	PRIOR APPLICATION NUMBER: 607/090948	
65	PRIOR FILING DATE: 1998-06-19	
66	PRIOR APPLICATION NUMBER: 607/089952	
67	PRIOR FILING DATE: 1998-06-19	
68	PRIOR APPLICATION NUMBER: 607/090254	
69	PRIOR FILING DATE: 1998-06-22	
70	PRIOR APPLICATION NUMBER: 607/090349	
71	PRIOR FILING DATE: 1998-06-23	
72	PRIOR APPLICATION NUMBER: 607/090355	
73	PRIOR FILING DATE: 1998-06-23	

1	PRIOR APPLICATION NUMBER: 60/090422
2	PRIOR FILING DATE: 1998-06-24
3	PRIOR APPLICATION NUMBER: 60/090431
4	PRIOR FILING DATE: 1998-06-24
5	PRIOR APPLICATION NUMBER: 60/090435
6	PRIOR FILING DATE: 1998-06-24
7	PRIOR APPLICATION NUMBER: 60/090444
8	PRIOR FILING DATE: 1998-06-24
9	PRIOR APPLICATION NUMBER: 60/090445
10	PRIOR FILING DATE: 1998-06-24
11	PRIOR APPLICATION NUMBER: 60/090472
12	PRIOR FILING DATE: 1998-06-24
13	PRIOR APPLICATION NUMBER: 60/090535
14	PRIOR FILING DATE: 1998-06-24
15	PRIOR APPLICATION NUMBER: 60/090540
16	PRIOR FILING DATE: 1998-06-24
17	PRIOR APPLICATION NUMBER: 60/090542
18	PRIOR FILING DATE: 1998-06-24
19	PRIOR APPLICATION NUMBER: 60/090557
20	PRIOR FILING DATE: 1998-06-24
21	PRIOR APPLICATION NUMBER: 60/090676
22	PRIOR FILING DATE: 1998-06-25
23	PRIOR APPLICATION NUMBER: 60/090678
24	PRIOR FILING DATE: 1998-06-25
25	PRIOR APPLICATION NUMBER: 60/090690
26	PRIOR FILING DATE: 1998-06-25
27	PRIOR APPLICATION NUMBER: 60/090694
28	PRIOR FILING DATE: 1998-06-25
29	PRIOR APPLICATION NUMBER: 60/090695
30	PRIOR FILING DATE: 1998-06-25
31	PRIOR APPLICATION NUMBER: 60/090696
32	PRIOR FILING DATE: 1998-06-25
33	PRIOR APPLICATION NUMBER: 60/090862
34	PRIOR FILING DATE: 1998-06-26
35	PRIOR APPLICATION NUMBER: 60/090863
36	PRIOR FILING DATE: 1998-06-26
37	PRIOR APPLICATION NUMBER: 60/091360
38	PRIOR FILING DATE: 1998-07-01
39	PRIOR APPLICATION NUMBER: 60/091478
40	PRIOR FILING DATE: 1998-07-02
41	PRIOR APPLICATION NUMBER: 60/091544
42	PRIOR FILING DATE: 1998-07-01
43	PRIOR APPLICATION NUMBER: 60/091519
44	PRIOR FILING DATE: 1998-07-02
45	PRIOR APPLICATION NUMBER: 60/091626
46	PRIOR FILING DATE: 1998-07-02
47	PRIOR APPLICATION NUMBER: 60/091633
48	PRIOR FILING DATE: 1998-07-02
49	PRIOR APPLICATION NUMBER: 60/091978
50	PRIOR FILING DATE: 1998-07-07
51	PRIOR APPLICATION NUMBER: 60/091982
52	PRIOR FILING DATE: 1998-07-07
53	PRIOR APPLICATION NUMBER: 60/092182
54	PRIOR FILING DATE: 1998-07-09

Query Match 10.9%; Score 155.5; DB 4; Length 358;
Best Local Similarity 47.5%; Pred. No. 1.6e-07;

[illegible]

RESULT 9
US-08-974-546-1
; Sequence 1, Application US/08974546
; Patent No. 5945287
; GENERAL INFORMATION:

```
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: TWO HUMAN HEAT SHOCK PROTEIN HOMOLOGS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,546
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0428
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAITUT21
; CLONE: 2525691
;
US-08-974-546-1

Query Match 10.7%; Score 154; DB 2; Length 348;
Best Local Similarity 45.0%; Pred. No. 2.1e-07;
Matches 27; Conservative 17; Mismatches 16; Indels 0; Gaps 0;

Qy 12 FHVIGVEATASDVVELKAYRQLAVMHPDKVHHPRAEAFKVLRAAWDIVSNAEKRYEYE 71
Db 6 YKILGIPSGANEDEIKKAYRKWALKYHPDKVKNKEPNAEKKFKEIAEAYDVLSDPKKRGLYD 65

RESULT 10
US-09-658-644-4
; Sequence 4, Application US/09658644
; Patent No. 6537554
; GENERAL INFORMATION:
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Jeffers, Michael
; TITLE OF INVENTION: Nucleotide Sequences and Amino Acid Sequences of
; FILE REFERENCE: 15966-517 CIP1
; CURRENT APPLICATION NUMBER: US/09/658,644
; CURRENT FILING DATE: 2001-09-09
; PRIOR FILING DATE: 1998-09-10
; PRIOR FILING DATE: 1998-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-658-644-4

Query Match 10.7%; Score 154; DB 2; Length 348;
Best Local Similarity 45.0%; Pred. No. 2.1e-07;
Matches 27; Conservative 17; Mismatches 16; Indels 0; Gaps 0;

Qy 12 FHVIGVEATASDVVELKAYRQLAVMHPDKVHHPRAEAFKVLRAAWDIVSNAEKRYEYE 71
Db 6 YKILGIPSGANEDEIKKAYRKWALKYHPDKVKNKEPNAEKKFKEIAEAYDVLSDPKKRGLYD 65

RESULT 10
US-09-658-644-4
; Sequence 4, Application US/09658644
; Patent No. 6537554
; GENERAL INFORMATION:
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Jeffers, Michael
; TITLE OF INVENTION: Nucleotide Sequences and Amino Acid Sequences of
; FILE REFERENCE: 15966-517 CIP1
; CURRENT APPLICATION NUMBER: US/09/658,644
; CURRENT FILING DATE: 2001-09-09
; PRIOR FILING DATE: 1998-09-10
; PRIOR FILING DATE: 1998-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-658-644-4

APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: TWO HUMAN HEAT SHOCK PROTEIN HOMOLOGS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,546
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0428
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAITUT21
CLONE: 2525691

US-08-974-546-1

Query Match 10.4%; Score 149.5; DB 4; Length 223;
Best Local Similarity 31.9%; Pred. No. 3.2e-07;
Matches 37; Conservative 25; Mismatches 37; Indels 17; Gaps 3;

Qy 12 FHVIGVEATASDVVELKAYRQLAVMHPDKVHHPRAEAFKVLRAAWDIVSNAEKRYEYE 71
Db 28 YDILGVKPSASERQIKKAFKFLAMKYPDKVSPDAEAFKEIAEAYETILSDANRRKEYD 87
Qy 72 M-----KMAENELSRSVN-EFLSKLODDLKEAMNTWMCSCQCGKRRFE 115
Db 88 TLGHSFTSGKGGSGSSFEQSFNF-----DDLFGFGFGQONQNTGSKKRF 138

RESULT 11
US-08-974-546-5
; Sequence 5, Application US/08974546
; Patent No. 5945287
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: TWO HUMAN HEAT SHOCK PROTEIN HOMOLOGS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,546
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0428
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 340 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1816452
;
US-08-974-546-5

Query Match 10.3%; Score 147; DB 2; Length 340;
Best Local Similarity 43.1%; Pred. No. 1.1e-06;
Matches 31; Conservative 14; Mismatches 25; Indels 2; Gaps 1;

Qy 12 FHVIGVEATASDVVELKAYRQLAVMHPDKVHHPRAEAFKVLRAAWDIVSNAEKRYEYE 71
Db 6 YQTLGLARGASDEIKRAYRRQALRVHPDKVKNKEPNAEKKFKEIAEAYDVLSDPKKREIFD 65
Qy 72 MKMAENELSR 83
Db 66 --RYEGGLKGS 75
```

```

RESULT 12
US-08-686-417-3      US-08-686-417-3
; Sequence 3, Application US/08686417
; Patent No. 5850018
; GENERAL INFORMATION:
; APPLICANT: Baszczyński, Chris
; APPLICANT: Barbour, Eric
; APPLICANT: Horowitz, Jeannine
; APPLICANT: Rosichan, Jeffrey L.
; TITLE OF INVENTION: AN EXPRESSION CONTROL SEQUENCE FOR
; TITLE OF INVENTION: GENERAL AND EFFECTIVE EXPRESSION OF GENES IN PLANTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/686,417
; FILING DATE: 26-JUL-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 27112-20038.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 822-0168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-686-417-3

Query Match          10.3%; Score 147; DB 2; Length 419;
Best Local Similarity 22.5%; Pred. No. 1.5e-06;
Matches 54; Conservative 37; Mismatches 67; Indels 82; Gaps 9;

QY    7  DELNPHVLGVETASDVELKKAYRQLAVMVHPDKNHPPAEAFKVLVRAADIVSNAEK 66
DB    10 DNTKYVEILGVPKSASDDLKAYRKAAIKNHDPKGDDP---EKFKELAQAYEVLSPEK 66
QY    67 RKEYEMKMAENEL-----SRSVNFIPLSK 90
DB    67 REIYD--QYGEDALKEGMGGGSHVDPFDFISFFPGSGGGSSRGRRQRGEDVVHP 124
QY    91 LODDLKEAMN-----TMMCSRCQGKHRRFMENDREPKSARYCAECNLRLHPAPEGDF 140
DB    125 LKVSLEDLYNGTSKKLSLRSNVICSKCKGK-----GSKSGASNRCPCG-----QGS- 170
QY    141 WABSSMLGLKITFYALMDKGKVYDITWAG-CQRVGISPDRHRVPYHISFGSRIPGTGRGQ 199
DB    171 -----GMKVTRIQLGPSMIQQMQPCNECKTGTS-----INCKDRCPGCKGEK 214

RESULT 13
US-09-149-476-497
; Sequence 497, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476

```



```
; APPLICANT: Sectrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370.838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 199
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-370-838-199

Query Match      10.0%; Score 144; DB 4; Length 127;
Best Local Similarity 46.9%; Pred. No. 5e-07;
Matches 30; Conservative 14; Mismatches 18; Indels 2; Gaps 1;

Qy      8 ELNPFHVLGVEATASDVLELKAYRQLAVMVHPDKNHHPRAEAEAFKVLRAAWDIVSNAEKR 67
      | : ||| : | : ||| ||| : ||| | : ||| : ||| : ||| : ||| : ||| : |||
Db      4 ETTYDVLGVKPNATQBELKKAYRKLALKYHPDKN--PNEGKFKQISQAYEVLSDAKKR 61

Qy      68 KEYE 71
      : :
Db      62 ELYD 65

RESULT 15
US-09-252-991A-27174
; Sequence 27174, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27174
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27174

Query Match      10.0%; Score 143.5; DB 4; Length 381;
Best Local Similarity 47.5%; Pred. No. 2.9e-06;
Matches 29; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

Qy      12 FHVILGVEATASDVLELKAYRQLAVMVHPDKNHHPR-AEEAFKVLRAAWDIVSNAEKEY 70
      : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||
Db      11 YEVILGVERGASADLKKAYRLAMKYHPDRNPGDKAEADFKKANEAYEVLSDAKRAAY 70

Qy      71 E 71
      :
Db      71 D 71

Search completed: January 2, 2004, 16:43:25
Job time : 26 secs
```



```
Query Match      100.0%; Score 1433; DB 12; Length 412;
Best Local Similarity 100.0%; Pred. No. 7.6e-127;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGVPEDELNPFHVLGVETASDELKAYRQLAVMHPDKNHHPRABEAFKVLRAAWDI 60
Db 144 MAGVPEDELNPFHVLGVETASDELKAYRQLAVMHPDKNHHPRABEAFKVLRAAWDI 203

Qy 61 VSNAEKRKEYEMKMAENELSRVNEFLSKLQDDLKEAMNTMCSRCCGKHRRFEMDREP 120
Db 204 VSNAEKRKEYEMKMAENELSRVNEFLSKLQDDLKEAMNTMCSRCCGKHRRFEMDREP 263

Qy 121 KSARYCAECNLHPAEEDGFWAESMLGLKITYFALMDGKYYDITWAGCORVGISPDTH 180
Db 264 KSARYCAECNLHPAEEDGFWAESMLGLKITYFALMDGKYYDITWAGCORVGISPDTH 323

Qy 181 RVPYHISFGSRIPGTRGRQRTAPDAPPADLQDFLSRIFQVPPGQMPNGNFFAAPQAPGA 240
Db 324 RVPYHISFGSRIPGTRGRQRTAPDAPPADLQDFLSRIFQVPPGQMPNGNFFAAPQAPGA 383

Qy 241 AAASKPNSTVPKGEAKPKRRKKVRRPFR 269
Db 384 AAASKPNSTVPKGEAKPKRRKKVRRPFR 412

RESULT 2
US-10-097-340-183
; Sequence 183, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAWATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLIS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHWANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 183
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-183
```

```
Query Match      58.7%; Score 841.5; DB 15; Length 191;
Best Local Similarity 88.1%; Pred. No. 2e-71;
Matches 163; Conservative 1; Mismatches 12; Indels 9; Gaps 3;

Qy 72 MKRMAENELSRVNEFLSKLQDDLKEAMNTMCSRCCGKHRRFEMDREPKSARYCAECNR 131
Db 1 MKRMAENELSRVNEFLSKLQDDLKEAMNTMCSRCCGKHRRFEMDREPKSARYCAECNR 60

Qy 132 LHPAEEDGFWAESMLGLKITYFALMDGKYYDITWAGCORVGISPDTHRVPYHISFGSR 191
Db 61 LHPAEEDGFWAESMLGLKITYFALMDGKYYDITWAGCORVGISPDTHRVPYHISFGSR 120

Qy 192 IPGTRGRQRTAPDAPPADLQDFLSRIFQVPPGQMPNG---NFFAAP---QAPAGAA--A 242
Db 121 IPGTRGRQRTAPDAPPADLQDFLSRIFQVPPGQMPNGTSLQLSLPLEPLQPLSPTTAQYP 180

Qy 243 ASKEN 247
Db 181 REKEN 185

RESULT 3
US-09-764-868-1161
; Sequence 1161, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ32
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1161
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (251)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-1161

Query Match      13.2%; Score 189; DB 10; Length 258;
Best Local Similarity 38.3%; Pred. No. 1.3e-09;
Matches 44; Conservative 14; Mismatches 27; Indels 30; Gaps 3;

Qy 12 FHVLGVEATASDELKAYRQLAVMHPDKNHHPRABEAFKVLRAAWDI VSNAEKRKEYE 71
Db 127 YEILGVSRGASDEDLKAYRRLLAKFHPDKNHAPGATEAFKAIGTAVAVLSNPEKRQYD 186

Qy 72 MKRMAENELSRVNEFLSKLQDDLKEAMNTMCSRCCGCK-----HRRFEMDREP 121
Db 187 -----QFGDDKSQAAR-----HGHGHGDFPHRGFEADISPE 216

RESULT 4
US-09-764-868-789
; Sequence 789, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ32
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 789
; LENGTH: 373
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-868-789

Query Match      12.5%; Score 179; DB 10; Length 373;
Best Local Similarity 47.4%; Pred. No. 1.9e-08;
Matches 36; Conservative 13; Mismatches 27; Indels 0; Gaps 0;

Qy 10 NPHVLGVEATASDVELKKAYRQLAVMVHPDKNHHPRAEAFKVLRAAWDIVSNAEKRYE 69
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 102 NYEVLGVTKDAGDEDLKKAYRKALKLFHDKHAPGATDAFKKIGNAYAVLSNPKRKQ 161
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 70 YEMKRWAENELSRSVN 85
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 162 YDLTGNEEQACNHQNN 177
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 5
US-09-764-868-1162
; Sequence 1162, Application US/09764869
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ32
; CURRENT APPLICATION NUMBER: US/09/764,868
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1162
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (110)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (143)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (193)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (268)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (288)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-1162

Query Match      12.0%; Score 172; DB 10; Length 297;
Best Local Similarity 46.1%; Pred. No. 6.4e-08;
Matches 35; Conservative 12; Mismatches 29; Indels 0; Gaps 0;

Qy 10 NPHVLGVEATASDVELKKAYRQLAVMVHPDKNHHPRAEAFKVLRAAWDIVSNAEKRYE 69
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 99 NYEVLGVTKDAGDEDLKKAYRKALKLFHDKHAPGATDAFKKIGNAYAVLSNPKRKQ 158
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 70 YEMKRWAENELSRSVN 85
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 159 YDLTGNEEQACNHQNN 174
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 6
US-10-369-493-134
; Sequence 134, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
```

```
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 134
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Aquifex aeolicus
US-10-369-493-134

Query Match      11.6%; Score 166.5; DB 12; Length 364;
Best Local Similarity 24.2%; Pred. No. 2.8e-07;
Matches 62; Conservative 40; Mismatches 97; Indels 57; Gaps 10;

Qy 12 FHVLGVEATASDVELKKAYRQLAVMVHPDKNHHPRAEAFKVLRAAWDIVSNAEKRYE 71
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 9 YEILGVNRDATKEEIKKAYRKLVRIYHPDINPDPSAQEKFEINEAYHVLIDDERSEYD 68
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 72 MKMAENELSRSVNEFLSKLQDDLKEAMNTMCSRCQKHRRFEMDREPKSARYCAECNR 131
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 69 -AILSRNDVGK-PRDFLEVIQEFVESIIQGE-----KGKKRRRKGDIDKMKPLT----- 117
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 132 LHPAEGDFWAESSMLGLKITYPALMDGKYYDITWAGQORVGISPDTHRVYHISFGSR 191
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 118 ---LEEAGLGCEK-----EIIYSRWMDCPV-----CEGMVGKEAETVWCHACNG-- 159
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 192 IPGTRGRQRATPDAPPADLIQDFLSRIFQVP-PCQPMNGHFFAAPQAPGAAAKPNSTV 250
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 160 ----EGRR-----VSGIFNFRPCSGCKGKGFVKNPCP-----TCY 192
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 251 PKGEAKPKRRKKVRRP 266
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 193 GRGRVSAQHKIKVHIP 208
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
US-09-764-868-1159
; Sequence 1159, Application US/09764869
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ32
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1159
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (206)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (212)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (213)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-1159

Query Match      11.5%; Score 165; DB 10; Length 216;
Best Local Similarity 33.8%; Pred. No. 1.9e-07;
Matches 45; Conservative 26; Mismatches 54; Indels 8; Gaps 4;
```

[illegible]

```

; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1163
; LENGTH: 132
; TYPE: PR
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (22)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (46)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (120)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (130)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-868-1163

Query Match 11.2%; Score 161; DB 10; Length 132;
Best Local Similarity 51.6%; Pred. No. 2.3e-07;
Matches 32; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

Qy 10 NPHVLGVGVEATSDVELKKAYRQLAVMVHPDKNHHPRAEAFKVLRAAWDIVSNAEKKE 69
Db 54 NYEILGVSRDASDEBELKKAYRKALKFHPDKNCAPGATDAFKAGNAFAVLSPDKRLR 113
Qy 70 YE 71
Db 114 YD 115

RESULT 12
US-10-161-051-134
; Sequence 134, Application US/10161051
; Publication No. US20030152945A1
; GENERAL INFORMATION:
; APPLICANT: Peter Deak
; APPLICANT: David M Glover
; APPLICANT: Carol Midgley
; TITLE OF INVENTION: Cell cycle progression proteins
; FILE REFERENCE: CCI-021CP
; CURRENT APPLICATION NUMBER: US/10/161,051
; CURRENT FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: GB 0007268.6
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 134
; LENGTH: 334
; TYPE: PR
; ORGANISM: Drosophila melanogaster
US-10-161-051-134

Query Match 11.1%; Score 159; DB 12; Length 334;
Best Local Similarity 48.3%; Pred. No. 1.3e-06;
Matches 29; Conservative 16; Mismatches 15; Indels 0; Gaps 0;

Qy 12 FHLVGVGVEATSDVELKKAYRQLAVMVHPDKNHHPRAEAFKVLRAAWDIVSNAEKKEYE 71
Db 6 YKILGLERKASDEIKKAYRKALKYHPDKNKSPOAEERFKEIAEAYEVLSDKKRIFD 65

RESULT 13
US-09-919-039-178
; Sequence 178, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US

```

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

1	PRIOR FILING DATE: 1998-07-02
2	PRIOR APPLICATION NUMBER: 60/091544
3	PRIOR FILING DATE: 1998-07-01
4	PRIOR APPLICATION NUMBER: 60/091519
5	PRIOR FILING DATE: 1998-07-02
6	PRIOR APPLICATION NUMBER: 60/091626
7	PRIOR FILING DATE: 1998-07-02
8	PRIOR APPLICATION NUMBER: 60/091633
9	PRIOR FILING DATE: 1998-07-02
10	PRIOR APPLICATION NUMBER: 60/091978
11	PRIOR FILING DATE: 1998-07-07
12	PRIOR APPLICATION NUMBER: 60/091982
13	PRIOR FILING DATE: 1998-07-07
14	PRIOR APPLICATION NUMBER: 60/092182
15	PRIOR FILING DATE: 1998-07-09

Query Match
10.9%; Score 155.5; DB 9; Length 358;

Best Local Similarity 47.5%; Pred. No. 3e-06;
Matches 29; Conservative 20; Mismatches 11; Indels 1; Gaps 1;

[illegible]

Qy 71 E 71

87 D 87

RESULT 15

US-09-989-723-148
; Sequence 148, Application US/09989723
; Patent No. US20020072092A1

; GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.
APPLICANT: Everett, Neal

APPLICANT: Ferrara, Napoleo

APPLICANT: Fong Sherman

APPLICANT: FONG, Siler M

APPLICANT: Gerritsen, Mary

APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J

APPLICANT: Grimaldi, J. Chri

; APPLICANT: Gurney, Austin L

APPLICANT: Kljavin, Ivar J.

APPLICANT: Napier, Mary A.
APPLICANT: Dan James

APPLICANT: Pail, James

APPLICANT: ROY, Margaret A.

APPLICANT: Stewart, Timothy

; APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin

APPLICANT: Williams, P. Mic

APPLICANT: Wood, William I
APPLICANT: Zhang Zong

APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted

;	TITLE OF INVENTION:	Acids
;	TITLE OF INVENTION:	Acids

FILE OF INVENTION: AC103
; FILE REFERENCE: P2730P1C62

; CURRENT APPLICATION NUMBER:

; CURRENT FILING DATE: 2001-

; PRIOR APPLICATION NUMBER: 60

; PRIOR FILING DATE: 1997-06-

; PRIOR APPLICATION NUMBER: 6
 ; PRIOR FILING DATE: 1987-10-

PRIOR FILING DATE: 1997-10-16
PRIOR APPLICATION NUMBER: 60/080,000

; PRIOR APPLICATION NUMBER: 6
 ; PRIOR FILING DATE: 1997-11-

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

PRIOR FILING DATE: 1997-11-11

10; PRIOR APPLICATION NUMBER: 60

; PRIOR FILING DATE: 1997-11-11

1	PRIOR APPLICATION NUMBER	60/075944
2	PRIOR FILING DATE	1998-02-25
3	PRIOR APPLICATION NUMBER	60/078910
4	PRIOR FILING DATE	1998-03-20
5	PRIOR APPLICATION NUMBER	60/083322
6	PRIOR FILING DATE	1998-04-28
7	PRIOR APPLICATION NUMBER	60/084600
8	PRIOR FILING DATE	1998-05-07
9	PRIOR APPLICATION NUMBER	60/087106
10	PRIOR FILING DATE	1998-05-28
11	PRIOR APPLICATION NUMBER	60/087607
12	PRIOR FILING DATE	1998-06-02
13	PRIOR APPLICATION NUMBER	60/087609
14	PRIOR FILING DATE	1998-06-02
15	PRIOR APPLICATION NUMBER	60/087759
16	PRIOR FILING DATE	1998-06-02
17	PRIOR APPLICATION NUMBER	60/087827
18	PRIOR FILING DATE	1998-06-03
19	PRIOR APPLICATION NUMBER	60/088020
20	PRIOR FILING DATE	1998-06-04
21	PRIOR APPLICATION NUMBER	60/088025
22	PRIOR FILING DATE	1998-06-04
23	PRIOR APPLICATION NUMBER	60/088026
24	PRIOR FILING DATE	1998-06-04
25	PRIOR APPLICATION NUMBER	60/088028
26	PRIOR FILING DATE	1998-06-04
27	PRIOR APPLICATION NUMBER	60/088029
28	PRIOR FILING DATE	1998-06-04
29	PRIOR APPLICATION NUMBER	60/088030
30	PRIOR FILING DATE	1998-06-04
31	PRIOR APPLICATION NUMBER	60/088033
32	PRIOR FILING DATE	1998-06-04
33	PRIOR APPLICATION NUMBER	60/088326
34	PRIOR FILING DATE	1998-06-04
35	PRIOR APPLICATION NUMBER	60/088167
36	PRIOR FILING DATE	1998-06-05
37	PRIOR APPLICATION NUMBER	60/088202
38	PRIOR FILING DATE	1998-06-05
39	PRIOR APPLICATION NUMBER	60/088212
40	PRIOR FILING DATE	1998-06-05
41	PRIOR APPLICATION NUMBER	60/088217
42	PRIOR FILING DATE	1998-06-05
43	PRIOR APPLICATION NUMBER	60/088655
44	PRIOR FILING DATE	1998-06-09
45	PRIOR APPLICATION NUMBER	60/088734
46	PRIOR FILING DATE	1998-06-10
47	PRIOR APPLICATION NUMBER	60/088738
48	PRIOR FILING DATE	1998-06-10
49	PRIOR APPLICATION NUMBER	60/088742
50	PRIOR FILING DATE	1998-06-10
51	PRIOR APPLICATION NUMBER	60/088910
52	PRIOR FILING DATE	1998-06-10
53	PRIOR APPLICATION NUMBER	60/088924
54	PRIOR FILING DATE	1998-06-10
55	PRIOR APPLICATION NUMBER	60/088926
56	PRIOR FILING DATE	1998-06-10
57	PRIOR APPLICATION NUMBER	60/088958
58	PRIOR FILING DATE	1998-06-11
59	PRIOR APPLICATION NUMBER	60/088961
60	PRIOR FILING DATE	1998-06-11
61	PRIOR APPLICATION NUMBER	60/088976
62	PRIOR FILING DATE	1998-06-11
63	PRIOR APPLICATION NUMBER	60/089105
64	PRIOR FILING DATE	1998-06-12
65	PRIOR APPLICATION NUMBER	60/089440
66	PRIOR FILING DATE	1998-06-16
67	PRIOR APPLICATION NUMBER	60/089512
68	PRIOR FILING DATE	1998-06-16
69	PRIOR APPLICATION NUMBER	60/089514
70	PRIOR FILING DATE	1998-06-16
71	PRIOR APPLICATION NUMBER	60/089532
72	PRIOR FILING DATE	1998-06-17
73	PRIOR APPLICATION NUMBER	60/089538

; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01

; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 10.9%; Score 155.5; DB 9; Length 358;
Best Local Similarity 47.5%; Pred. No. 3e-06;
Matches 29; Conservative 20; Mismatches 11; Indels 1; Gaps 1;

Qy 12 FHVIGVEATASDVLEKAYEQQLAVMGHPDQV-HHPRAEEAFKVLRAAWDIVSNAEKKEY 70
Db 27 YKILGVPRASIKDIKAYEKALQLHPRDRNPDPOAQEKFDLGAAYEVLSDSEKQY 86
Qy 71 E 71
Db 87 D 87

Search completed: January 2, 2004, 16:44:33
Job time : 56 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 16:41:29 ; Search time 42 Seconds

(without alignments)
1016.606 Million cell updates/sec

Title: US-10-049-742-11

Perfect score: 269

Sequence: 1 MAGVPEDELNPFHVLGVEAT.....VPKGEAKPKRKKVRRPFQR 269

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Listing first 45 summaries

Database : A Geneseq 19Jun03.*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	269	100.0	269	22	AAB67455 Amino acid sequenc
2	242	90.0	702	23	ABY77732 Amino acid sequenc
3	168	62.5	412	23	ABY77019 Human protein sequ
4	168	62.5	412	23	ABG64927 Human albumin fusi
5	155	57.6	191	20	AA32126 Human LysT interac
6	155	57.6	191	20	ABG96364 Human ovarian carc
7	95	35.3	699	23	AA371490 Bovine Jiv protein
8	90	33.5	90	23	AA371492 Human Jiv protein
9	89	33.1	90	23	AA371491 Bovine Jiv protein

10	81	30.1	3835	22	AAB31167	Amino acid sequenc
11	81	30.1	3988	21	AAV53615	NADL protein encod
12	81	30.1	3988	21	AAV53616	Amino acid sequenc
13	53	19.7	204	23	ABY77045	Human protein sequ
14	53	19.7	204	23	ABG64926	Human albumin fusi
15	14	5.2	970	22	ABG69643	Drosophila melanog
16	8	3.0	155	22	ABY11029	Human SPR2B homolo
17	8	3.0	155	22	AAW79669	Human protein SEQ
18	8	3.0	182	24	ABF56862	Mouse CITEF4 prote
19	8	3.0	186	24	ABU02261	S. pneumoniae type
20	8	3.0	295	20	AAV23861	A Trichomonas vagi
21	8	3.0	297	20	AAV23860	A Trichomonas vagi
22	8	3.0	299	20	AAV23859	A Trichomonas vagi
23	8	3.0	607	23	AA666030	Amino acid sequenc
24	8	3.0	1285	21	AA843949	Human cancer assoc
25	8	3.0	1451	23	AAU81018	Human alpha2 macro
26	8	3.0	1474	12	AAV11334	Recombinant human
27	8	3.0	1474	21	AAV97157	Human alpha-2-macr
28	8	3.0	1474	22	AA850673	Human alpha-2-macr
29	8	3.0	1474	23	ABF65218	Hypoxia-regulated
30	8	3.0	1474	23	AAU74798	Human alpha 2 macr
31	8	3.0	1474	23	AAU81017	Human alpha2 macro
32	8	3.0	1484	12	AAV11749	Human alpha-2 macr
33	8	3.0	1714	22	AB860186	Drosophila melanog
34	7	2.6	46	23	AA666029	Signal peptide for
35	7	2.6	62	22	ABG55351	Human liver peptid
36	7	2.6	62	22	AB840166	Peptide #7672 enco
37	7	2.6	62	22	AB824617	Protein #6616 enco
38	7	2.6	62	22	AAV60932	Human brain expres
39	7	2.6	62	22	AAV73618	Human bone marrow
40	7	2.6	62	22	AAV20041	Peptide #6475 enco
41	7	2.6	62	22	AAV33813	Peptide #7850 enco
42	7	2.6	62	23	AAV43488	Human peptide enco
43	7	2.6	75	21	AAV01935	Human secreted pro
44	7	2.6	101	19	AAV85839	S. pneumoniae deri
45	7	2.6	102	21	AAV01056	Human secreted pro

ALIGNMENTS

RESULT 1

AAB67455

ID AAB67455 standard; Protein; 269 AA.

XX AAB67455;

XX 15-MAY-2001 (first entry)

XX Amino acid sequence of a human chaperone polypeptide.

XX Human; chaperone polypeptide; reproductive disease; prolactin production;
XX infertility; tumour; cancer; Peyronie's disease; eye disorder; glaucoma;
XX conjunctivitis; keratitis; neuromuscular disorder; cystic fibrosis;
XX metabolic disorder; Zellweger syndrome; Addison's disease; iritis;
XX autoimmune disorder; inflammatory disorder; systemic lupus erythematosus;
XX acquired immunodeficiency syndrome; AIDS; asthma; atherosclerosis;
XX cell proliferative disorder; gene therapy.

XX Homo sapiens.

XX W0200109178-A2.

XX 08-FEB-2001.

XX 03-AUG-2000; 2000WO-US21313.

XX 03-AUG-1999; 99US-0146908.

XX 22-OCT-1999; 99US-0160924.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Bandman O, Tang YT, Baughn MR, Azimzai Y, Lu DAM;

XX WPI; 2001-159853/16.
DR N-PSDB; AAF54994.
XX
PT New human chaperone proteins and polynucleotides, useful in diagnosing,
PT treating and preventing reproductive, eye, neuromuscular, metabolic,
PT autoimmune or inflammatory disorders -
XX
PS Claim 1; Page 101-102; 102pp; English.
XX
CC The present sequence represents a human chaperone polypeptide. Human
CC chaperone polypeptides and polynucleotides are useful in the diagnosis,
CC treatment and prevention of reproductive (e.g. prolactin production,
CC infertility, endometrial or ovarian tumour, cancer of the breast,
CC prostate or testis, Peyronie's disease), eye (e.g. conjunctivitis,
CC keratitis, iritis, glaucoma), neuromuscular, metabolic (e.g. Zellweger
CC syndrome, Addison's disease, cystic fibrosis), and autoimmune and
CC inflammatory disorders (e.g. systemic lupus erythematosus, acquired
CC immunodeficiency syndrome (AIDS), asthma, atherosclerosis), infectious
CC or viral diseases, and cell proliferative disorders. Chaperone
CC polynucleotides may be used for somatic or germline gene therapy, to
CC detect and quantify gene expression in biopsied tissues in which
CC expression is correlated with disease.
XX
SQ Sequence 269 AA;
Query Match 100.0%; Score 269; DB 22; Length 269;
Best Local Similarity 100.0%; Pred. No. 2.9e-248;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAGVPDELNPFHVLGVEATSDVELKAYRQLAVMVHPDKNHHPRAEAFKVLRAAWDI 60
Db 1 MAGVPDELNPFHVLGVEATSDVELKAYRQLAVMVHPDKNHHPRAEAFKVLRAAWDI 60
Qy 61 VSNAEKKEYEMKMAENELSRVNEFLSKLQDDLKEAMNTMCSRCQGHRRFEMDREP 120
Db 61 VSNAEKKEYEMKMAENELSRVNEFLSKLQDDLKEAMNTMCSRCQGHRRFEMDREP 120
Qy 121 KSARYCAECNRLHPAEEGDFWAESSMLGLKITTYFALMDGKVDITWAGCQGVISPDTH 180
Db 121 KSARYCAECNRLHPAEEGDFWAESSMLGLKITTYFALMDGKVDITWAGCQGVISPDTH 180
Qy 181 RVPYHISFGSRIPGTRGQRATPDAPPADLQDLFLSRIFQVPPGMPNGNFFAAPPAPGA 240
Db 181 RVPYHISFGSRIPGTRGQRATPDAPPADLQDLFLSRIFQVPPGMPNGNFFAAPPAPGA 240
Qy 241 AAASKPNSTVPKGEAKPKRRKKVRRPFR 269
Db 241 AAASKPNSTVPKGEAKPKRRKKVRRPFR 269
RESULT 2
ABB77732
ID ABB77732 standard; Protein; 702 AA.
XX
AC ABB77732;
XX
DT 30-JUL-2002 (first entry)
XX
DE Amino acid sequence of human P125-77.22 polypeptide.
XX
KW Human; P125-77.22; mucosal disease; BVDV infection; gene therapy.
XX
OS Homo sapiens.
XX
XX WO200226810-A1.
XX
PD 04-APR-2002.
XX
XX 10-SEP-2001; 2001WO-CN01354.
PF
XX
PR 12-SEP-2000; 2000CN-0125190.
XX

PA (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.
XX
XX Mao Y, Xie Y;
XX
XX WPI; 2002-281319/32.
DR N-PSDB; ABL56700.
XX
PT Protein P125-77.22 and encoding polynucleotide, used in diagnosis and
PT treatment of human mucosal disease caused by BVDV infection -
XX
PS Claim 1; Page 29-30; 33pp; Chinese.
XX
CC The present sequence represents human P125-77.22 polypeptide. The
CC polypeptide and polynucleotide are used in diagnosis and treatment
CC of human mucosal disease caused by viral BVDV (undefined) infection.
CC The polynucleotide may also be used for gene therapy.
XX
SQ Sequence 702 AA;
Query Match 90.0%; Score 242; DB 23; Length 702;
Best Local Similarity 100.0%; Pred. No. 3.8e-222;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 28 KAYRQLAVMVHPDKNHHPRAEAFKVLRAAWDIVSNAEKKEYEMKMAENELSRVNEF 87
Db 461 KAYRQLAVMVHPDKNHHPRAEAFKVLRAAWDIVSNAEKKEYEMKMAENELSRVNEF 520
Qy 88 LSKLQDDLKEAMNTMCSRCQGHRRFEMDREPDKSARYCAECNRLHPAEEGDFWAESSML 147
Db 521 LSKLQDDLKEAMNTMCSRCQGHRRFEMDREPDKSARYCAECNRLHPAEEGDFWAESSML 580
Qy 148 GLKITTYFALMDGKVDITWAGCQGVISPDTHRVPYHISFGSRIPGTRGQRATPDAPP 207
Db 581 GLKITTYFALMDGKVDITWAGCQGVISPDTHRVPYHISFGSRIPGTRGQRATPDAPP 640
Qy 208 ADLQDLFLSRIFQVPPGMPNGNFFAAPPAPGAASAKPNSTVPKGEAKPKRRKKVRRPF 267
Db 641 ADLQDLFLSRIFQVPPGMPNGNFFAAPPAPGAASAKPNSTVPKGEAKPKRRKKVRRPF 700
Qy 269 QR 269
Db 701 QR 702
RESULT 3
ABB777019
ID ABB777019 standard; Protein; 412 AA.
XX
XX ABB777019;
XX
DT 08-OCT-2002 (first entry)
XX
DE Human protein sequence #1 from clone HNTPB82.
XX
KW Human; HNTPB82; secreted protein; immunosuppressive; food preservative;
KW antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;
KW vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;
KW virucide; fungicide; ophthalmological; vulnerary; gene therapy; ELISA;
KW radioimmunoassay; enzyme linked immunosorbent assay; autoimmune disease;
KW hyperproliferative disorder; cardiovascular disorder; angiogenesis;
KW cerebrovascular disorder; nervous system disorder; ocular disorder;
KW wound healing; food additive.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT Peptide 1 /label= Signal_peptide
FT Protein 2..412 /label= Mature_protein
XX
XX WO200222638-A1.
XX

PD 21-MAR-2002.
XX
XX 17-JAN-2001; 2001WO-US01386.
XX
XX 12-SEP-2000; 2000US-232104P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;
PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;
PI Ni J;
XX
XX WPI; 2002-258041/30.
DR N-PSDB; ABL55088.
XX
XX New nucleic acid molecules encoding 22 human secreted proteins for
PT diagnosing or treating e.g. autoimmune diseases, hyperproliferative
PT disorders, and cardiovascular disorders, and used as food additives or
PT preservatives -
XX
XX Disclosure; Page 486-487; 526pp; English.
XX
XX The sequence represents a protein sequence of the invention, encoded by
CC cDNA isolated from human clone ID HNT882. The invention relates to novel
CC isolated nucleic acid molecules encoding 22 human secreted proteins. The
CC proteins of the invention have immunosuppressive, antiarthritic,
CC antirheumatic, antiproliferative, cytostatic, cardiant, vasotropic,
CC cerebroprotective, neurotropic, neuroprotective, antibacterial, virucide,
CC fungicide, ophthalmological, and vulnery activity. The polynucleotides
CC may have a use in gene therapy. The polynucleotides and polypeptides
CC encoded by them are used to prevent, treat or ameliorate a medical
CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
CC chickens or sheep. The polynucleotides and polypeptides are also used in
CC diagnosing a pathological condition or susceptibility to a pathological
CC condition. The antibodies to the proteins can also be used in alleviating
CC symptoms associated with the disorders and in diagnostic immunoassays
CC e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA).
CC Disorders which are diagnosed or treated include autoimmune diseases,
CC hyperproliferative disorders, cardiovascular disorders, cerebrovascular
CC disorders, angiogenesis, nervous system disorders, infections caused by
CC bacteria, viruses and fungi and ocular disorders. The polypeptides can
CC also be used to aid wound healing and epithelial cell proliferation. The
CC polypeptides can also be used as a food additive or preservative.
XX
SQ Sequence 412 AA;
Query Match 62.5%; Score 168; DB 23; Length 412;
Best Local Similarity 99.6%; Pred. No. 9.8e-152;
Matches 268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAGVPEDELNPFHVLGVEATASDVLEKKAQRLAVMVHPDKNHHPRAEAFKVLRAAWDI 60
DB 144 MAGVPEDELNPFHVLGVEATASDVLEKKAQRLAVMVHPDKNHHPRAEAFKVLRAAWDI 203
QY 61 VSNAEKKEKYEKMAENELSRVNEFLSKLQDDLKEAMNTMCSRQCGKHRRFEMDREP 120
DB 204 VSNAEKKEKYEKMAENELSRVNEFLSKLQDDLKEAMNTMCSRQCGKHRRFEMDREP 263
QY 121 KSARYCAECNRLHPAEGDFWAESSMLGLKITYPALMDGKVYDITWAGCQRVIGISPDTH 180
DB 264 KSARYCAECNRLHPAEGDFWAESSMLGLKITYPALMDGKVYDITWAGCQRVIGISPDTH 323
QY 181 RVPYHISFGSRIPTGRGRQATPDAPPADLQDFLSRIFQVPPGMPNGNFAAPQAPGA 240
DB 324 RVPYHISFGSRIPTGRGRQATPDAPPADLQDFLSRIFQVPPGMPNGNFAAPQAPGA 383
QY 241 AAASKPNSTVPKGAEPKRRKVRPQR 269
DB 384 AAASKPNSTVPKGAEPKRRKVRPQR 412
RESULT 4
ABG64927

ID ABG64927 standard; Protein; 412 AA.
XX
XX AC ABG64927;
XX
XX DT 27-AUG-2002 (first entry)
XX
XX DE Human albumin fusion protein #1602.
XX
XX KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW haematopoietic disorder; neural disorder; connective disorder;
KW cytostatic; antiinfertility; antiinflammatory; antiulcer;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; neurotropic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.
XX
XX OS Homo sapiens.
XX
XX OS Synthetic.
XX
XX PN WO200177137-A1.
XX
XX PN 18-OCT-2001.
XX
XX PD 12-APR-2001; 2001WO-US11988.
XX
XX PF 12-APR-2000; 2000US-229358P.
XX
XX PR 25-APR-2000; 2000US-199384P.
XX
XX PR 21-DEC-2000; 2000US-256931P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX PA Rosen CA, Haseltine WA;
XX
XX PI WPI; 2002-010886/01.
XX
XX DR New fusion protein for treating disease e.g. diabetes comprises an
XX PT albumin fused to a therapeutic protein -
XX
XX PS Claim 1; Page 1606-1607; 2102pp; English.
XX
XX CC The present invention relates to albumin fusion proteins comprising a
XX CC therapeutic protein X and human albumin (HA, also known as human serum
XX CC albumin, HSA). The proteins are useful for treating a disease or
XX CC disorder that may be modulated by therapeutic protein X. The albumin
XX CC extends the shelf-life of protein X, and may increase its biological
XX CC in vitro/in vivo activity. The protein is useful for treating and
XX CC diagnosing disorders such as cancer, reproductive disorders, digestive
XX CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
XX CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
XX CC (e.g. diabetes), haematopoietic disorders, neural disorders
XX CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
XX CC encephalomyelitis, meningitis, schizophrenia), and connective disorders
XX CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin
XX CC fusion proteins of the invention.
XX
SQ Sequence 412 AA;
Query Match 62.5%; Score 168; DB 23; Length 412;
Best Local Similarity 99.6%; Pred. No. 9.8e-152;
Matches 268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAGVPEDELNPFHVLGVEATASDVLEKKAQRLAVMVHPDKNHHPRAEAFKVLRAAWDI 60
DB 144 MAGVPEDELNPFHVLGVEATASDVLEKKAQRLAVMVHPDKNHHPRAEAFKVLRAAWDI 203
QY 61 VSNAEKKEKYEKMAENELSRVNEFLSKLQDDLKEAMNTMCSRQCGKHRRFEMDREP 120
DB 204 VSNAEKKEKYEKMAENELSRVNEFLSKLQDDLKEAMNTMCSRQCGKHRRFEMDREP 263
QY 121 KSARYCAECNRLHPAEGDFWAESSMLGLKITYPALMDGKVYDITWAGCQRVIGISPDTH 180
DB 264 KSARYCAECNRLHPAEGDFWAESSMLGLKITYPALMDGKVYDITWAGCQRVIGISPDTH 323

QY 181 RVPYHISFGSRIPGTRGRQATPDAPPADLQDFLSRIFQVPPGQMPNGNFFFAAPQAPGA 240
DB 324 RVPYHISFGSRIPGTRGRQATPDAPPADLQDFLSRIFQVPPGQMPNGNFFFAAPQAPGA 383
QY 241 AAASKNSTVPKGEAKPKRKKVRRPFQR 269
DB 384 AAASKNSTVPKGEAKPKRKKVRRPFQR 412

RESULT 5
AAI32126
ID AAY32126 standard; Protein; 191 AA.
XX AC AAY32126;
XX DT 01-FEB-2000 (first entry)
XX DE Human LYST interacting protein LIP6.
XX KW LIP6; human; LYST-2; LYST interacting protein; lysosome; CHS;
KW Chediak-Higashi syndrome; autoimmune disease; rheumatoid arthritis;
KW systemic lupus erythematosus; inflammatory bowel disease;
KW diabetes mellitus; multiple sclerosis; atopic disease; asthma;
KW hay fever; rhinitis; urticaria; nasal polyp; cancer;
KW neurodegenerative disease; pigmentation disorder; viral disease;
KW platelet dysfunction.
XX KW Homo sapiens.
XX PN WO9951741-A2.
XX PD 14-OCT-1999.
XX PF 29-MAR-1999; 99WO-US06831.
XX PR 03-APR-1998; 98US-0054956.
XX PA (CURA-) CURAGEN CORP.
XX PI Nandabalan K, Kingsmore S;
XX WPI; 1999-620203/53.
XX N-PSDB; AAZ34492.
XX Protein complexes, interacting proteins, and related polynucleotides
XX useful for treating and preventing e.g. atopic, autoimmune or
XX neurodegenerative diseases -
XX Claim 20; Fig 7; 172pp; English.

XX The present sequence represents a novel human LYST interacting
XX protein, LIP6, that shows homology to pestivirus type 3 NS2-3.
XX LYST is the human lysosomal Chediak-Higashi syndrome (CHS) protein.
XX The invention relates to complexes of LYST or LYST-2 (see AAY32120)
XX with proteins identified as interacting with LYST or LYST-2 by a
XX modified yeast two-hybrid assay system. The interacting proteins
XX include 10 novel proteins, LIP1-10 (see AAY32121-30). Methods of
XX screening the protein complexes for efficacy in treating and/or
XX preventing atopic diseases (e.g. asthma, nasal polyps, hay fever
XX rhinitis, urticaria) autoimmune diseases (e.g. CHS, rheumatoid
XX arthritis, systemic lupus erythematosus, inflammatory bowel disease,
XX certain cancers, pigmentation disorders, platelet dysfunction and
XX viral diseases are provided. Nucleic acids (see AAZ34487-96)
XX encoding LIP1-10, modulation of LIP function by gene therapy, use
XX of antisense oligonucleotides for suppression of LIP protein
XX expression, screening for agonists and antagonists, diagnosing or
XX screening for the presence of a predisposition to a disease or
XX disorder, and animal models are also disclosed.

Sequence 191 AA;

Query Match 57.6%; Score 155; DB 20; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.3e-139;
Matches 155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 72 MKRMAENELSRVNEFLSKLQDDLKEAMNTMNCSCQGHRRFEMDRPKSARYCAECNR 131
DB 1 MKRMAENELSRVNEFLSKLQDDLKEAMNTMNCSCQGHRRFEMDRPKSARYCAECNR 60
QY 132 LHPAEGDFWABSSMLGLKITFYFALMDGKVYDITWAGCORVGISPDTHRVPHISFGSR 191
DB 61 LHPAEGDFWABSSMLGLKITFYFALMDGKVYDITWAGCORVGISPDTHRVPHISFGSR 120
QY 192 IPGTRGRQATPDAPPADLQDFLSRIFQVPPGQMP 226
DB 121 IPGTRGRQATPDAPPADLQDFLSRIFQVPPGQMP 155

RESULT 6
ABG96364
ID ABG96364 standard; Protein; 191 AA.
XX AC ABG96364;
XX DT 11-DEC-2002 (first entry)
XX DE Human ovarian cancer marker OV38.
XX KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;
KW central nervous system disorder; bacterial meningitis; viral meningitis;
KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
KW brain herniation; inflammation; encephalitis; testicular disorder;
KW nontuberculous granulomatous orchitis; connective tissue disorder;
KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
KW histological type; carcinogenic; ovarian cancer marker.
XX KW Homo sapiens.
XX PN WO200271928-A2.
XX PD 19-SEP-2002.
XX PF 14-MAR-2002; 2002WO-US07826.
XX PR 14-MAR-2001; 2001US-276025P.
XX PR 14-MAR-2001; 2001US-276026P.
XX PR 10-AUG-2001; 2001US-311732P.
XX PR 19-SEP-2001; 2001US-323580P.
XX PR 26-SEP-2001; 2001US-324967P.
XX PR 26-SEP-2001; 2001US-325102P.
XX PR 26-SEP-2001; 2001US-325149P.

XX (MILL-) MILLENNIUM PHARM INC.
XX Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
XX Meyers RE, Morrissey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
XX Baat RC, Lu K, Schmandt RE, Zhao X, Glatt K;
XX WPI; 2002-723277/78.
XX N-PSDB; ABS76460.
XX Assessing whether a patient is afflicted with ovarian cancer, useful in
XX assessing the stage or progression of the disease, comprises comparing
XX the expression level of a cancer marker in a sample from a patient and
XX from a non cancer patient -
XX Disclosure; Page 300; 481pp; English.

XX The present invention relates to a new method for assessing whether a
XX patient is afflicted with ovarian cancer. The method involves comparing
XX the expression level of a marker in a patient sample and the normal level
XX of expression of the marker in a control non-ovarian cancer sample, where
XX the marker is selected from 363 cancer markers described in the
XX specification. The method of the invention is useful in diagnosing or
XX

CC characterising cancer, in detecting the presence of cancer as early as
CC possible, and the recurrence of ovarian cancer. The method may also be of
CC particular use with patients having an enhanced risk of developing
CC ovarian cancer (e.g. patients having a familial history of ovarian
CC cancer). The cancer markers may be used in the management and treatment
CC of e.g. brain and central nervous system disorders (e.g. bacterial and
CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
CC inflammations (e.g. bacterial or viral meningitis or encephalitis),
CC testicular disorders (e.g. nontuberculous granulomatous orchitis),
CC connective tissue disorders, or heart disorders (e.g. ischaemic heart
CC disease or atherosclerosis). The compositions and methods may also be
CC used in assessing the histological type of neoplasm associated with
CC ovarian cancer, monitoring the progression of ovarian cancer,
CC determining whether ovarian cancer has metastasized or is likely to
CC metastasize, selecting a composition for inhibiting ovarian cancer,
CC assessing the ovarian carcinogenic potential of a compound, or
CC inhibiting ovarian cancer or at risk of developing ovarian cancer. The
CC present amino acid sequence represents one of the ovarian cancer markers
CC described in the invention.

XX Sequence 191 AA;
SQ Query Match 57.6%; Score 155; DB 23; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.3e-139;
Matches 155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 MKMAENELSRVNEFLSKLQDLKZAMNTMCSRQCGKHRRFEMDREPKSARYCAECNR 131
DB 1 MKMAENELSRVNEFLSKLQDLKZAMNTMCSRQCGKHRRFEMDREPKSARYCAECNR 60

QY 132 LHPAEEGDFWAESSMLGLKITYFALMDGKYVDITWAGCQGVGSPDTHRVPHYHISFGSR 191
DB 61 LHPAEEGDFWAESSMLGLKITYFALMDGKYVDITWAGCQGVGSPDTHRVPHYHISFGSR 120

QY 192 IPOTGRQRATPDAPPADLQDFLSRIFQVPPGOMP 226
DB 121 IPOTGRQRATPDAPPADLQDFLSRIFQVPPGOMP 155

RESULT 7
AAB71490
ID AAB71490 standard; protein; 699 AA.
AC AAB71490;
XX
XX
XX 28-NOV-2002 (first entry)
XX Bovine Jiv protein.
XX NS2; non-structural protein 2; inhibitor; competitive; allosteric; Jiv;
XX J-domain protein interacting with viral protein; NS2-3 viral protease;
XX virucide; hepatotropic; antiinflammatory; infection; bovine.
XX
XX Bos taurus.
XX
XX DE10112748-A1.
XX
XX 19-SEP-2002.
XX 14-MAR-2001; 2001DE-1012748.
XX 14-MAR-2001; 2001DE-1012748.
XX (TRAN-) TRANSMIT GES TECHNOLOGIETRANSFER MBH.
XX Tautz N, Thiel H, Birghan C;
XX WPI; 2002-692596/75.
XX New binding partners for hepatitis C virus non-structural protein 2,
XX useful for diagnosis, prevention and treatment of hepatitis C infection
PT -

XX Disclosure; Fig 2A; 14pp; German.
XX This invention describes a novel binding partner for non-structural
CC protein 2 (NS2) of hepatitis C virus (HCV) that competitively or
CC allosterically inhibits binding of Jiv (J-domain protein interacting
CC with viral protein) to NS2. Inhibiting Jiv and NS2 binding prevents
CC stimulation of NS2-3 viral protease. The products of the invention have
CC virucide, hepatotropic and antiinflammatory activity. The novel
CC binding partner is used for diagnosis, prevention and treatment of HCV
CC infection. This sequence represents the bovine Jiv protein described in
CC the disclosure of the invention.

XX Sequence 699 AA;
SQ Query Match 35.3%; Score 95; DB 23; Length 699;
Best Local Similarity 100.0%; Pred. No. 7e-82;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 EAMNTMCSRQCGKHRRFEMDREPKSARYCAECNRLHPAEEGDFWAESSMLGLKITYFAL 156
DB 527 EAMNTMCSRQCGKHRRFEMDREPKSARYCAECNRLHPAEEGDFWAESSMLGLKITYFAL 586

QY 157 MDGKYVDITWAGCQGVGSPDTHRVPHYHISFGSR 191
DB 587 MDGKYVDITWAGCQGVGSPDTHRVPHYHISFGSR 621

RESULT 8
AAB71492
ID AAB71492 standard; protein; 90 AA.
XX AAB71492;
XX 28-NOV-2002 (first entry)
XX Human Jiv protein Jiv90 fragment.
XX NS2; non-structural protein 2; inhibitor; competitive; allosteric; Jiv;
XX J-domain protein interacting with viral protein; NS2-3 viral protease;
XX virucide; hepatotropic; antiinflammatory; infection; human; Jiv90.
XX Homo sapiens.
XX DE10112748-A1.
XX 19-SEP-2002.
XX 14-MAR-2001; 2001DE-1012748.
XX 14-MAR-2001; 2001DE-1012748.
XX (TRAN-) TRANSMIT GES TECHNOLOGIETRANSFER MBH.
XX Tautz N, Thiel H, Birghan C;
XX WPI; 2002-692596/75.
XX New binding partners for hepatitis C virus non-structural protein 2,
XX useful for diagnosis, prevention and treatment of hepatitis C infection
PT -
XX Disclosure; Fig 3; 14pp; German.
XX This invention describes a novel binding partner for non-structural
CC protein 2 (NS2) of hepatitis C virus (HCV) that competitively or
CC allosterically inhibits binding of Jiv (J-domain protein interacting
CC with viral protein) to NS2. Inhibiting Jiv and NS2 binding prevents
CC stimulation of NS2-3 viral protease. The products of the invention have
CC virucide, hepatotropic and antiinflammatory activity. The novel
CC binding partner is used for diagnosis, prevention and treatment of HCV
CC infection. This sequence represents the human Jiv protein Jiv90 fragment
CC described in the disclosure of the invention.

63 DITEWAGCORVGISPDTHRVPYHISFGSR 191

RESULT 11
AA53615
ID AA53615 standard; Protein; 3988 AA.
XX
AC AA53615;
XX
DT 11-FEB-2000 (first entry)
XX
DE NADL protein encoded by the low copy number plasmid pACNR/BVDV NADL.
XX
KW Pseudorevertant; RNA virus; chimera; BVDV; HCV; replication-competent;
KW 5' nontranslated region; 5'NTR; 3' NTR; pestivirus; antiviral;
KW bovine viral diarrhea virus; NADL; vaccine.
XX
OS Synthetic.
OS Bovine viral diarrhea virus.
XX
PN WO9955366-A1.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-US08850.
XX
PR 24-APR-1998; 98US-0082964.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Rice CM, Frolov I, McBride MS;
XX
DR WPI; 2000-013359/01.
DR N-PSDB; AA236195.
XX
PT Chimeric viral RNA, used in vaccine against BVDV -
XX
PS Disclosure; Fig 10; 108pp; English.
XX
CC The present sequence represents the NADL protein of bovine viral diarrhea
CC virus (BVDV), and is encoded by the low copy number plasmid pACNR/BVDV
CC NADL. The plasmid is used in the course of the invention, to produce
CC chimeric RNA viruses. The specification describes chimeric viral
CC RNA comprising a 5' nontranslated region (5'NTR); an open reading frame
CC (ORF) region; and a 3' NTR; where at least one of the regions is chimeric
CC with a heterologous nucleotide sequence from a pestivirus in operable linkage
CC chimeric viral RNA is replication-competent. The chimeric viral RNA
CC can be used in a method for identifying compounds having antiviral
CC activity against HCV. When the pestivirus viral nucleotide sequence is
CC from bovine viral diarrhea virus (BVDV), the chimeric viral RNA can be
CC used in a vaccine against BVDV.
XX
SQ Sequence 3988 AA;
Query Match 30.1%; Score 81; DB 21; Length 3988;
Best Local Similarity 100.0%; Pred. No. 6.9e-68;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 103 MCSRCQKRRFFEMDRPKSARYCAECNRLHPAEEDGFWAESSMLGLKITFYALMDGKVY 162
DB 1537 MCSRCQKRRFFEMDRPKSARYCAECNRLHPAEEDGFWAESSMLGLKITFYALMDGKVY 1596
QY 163 DITEWAGCQRVGISPDTHRPV 183
DB 1597 DITEWAGCQRVGISPDTHRPV 1617
RESULT 12
AA53616
ID AA53616 standard; Protein; 3988 AA.
XX
AC AA53616;
XX
DT 11-FEB-2000 (first entry)

XX Amino acid sequence of infectious BVDV NADL protein.
XX
KW Pseudorevertant; RNA virus; chimera; BVDV; HCV; replication-competent;
KW 5' nontranslated region; 5'NTR; 3' NTR; pestivirus; antiviral;
KW bovine viral diarrhea virus; NADL; vaccine.
XX
OS Bovine viral diarrhea virus.
XX
PN WO9955366-A1.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-US08850.
XX
PR 24-APR-1998; 98US-0082964.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Rice CM, Frolov I, McBride MS;
XX
DR WPI; 2000-013359/01.
DR N-PSDB; AA239596.
XX
PT Chimeric viral RNA, used in vaccine against BVDV -
XX
PS Disclosure; Fig 11; 108pp; English.
XX
CC The present sequence represents the NADL protein of bovine viral diarrhea
CC virus (BVDV). The sequence is used in the course of the invention, to
CC produce chimeric RNA viruses. The specification describes chimeric viral
CC RNA comprising a 5' nontranslated region (5'NTR); an open reading frame
CC (ORF) region; and a 3' NTR; where at least one of the regions is chimeric
CC with a heterologous nucleotide sequence from a pestivirus in operable linkage
CC chimeric viral RNA is replication-competent. The chimeric viral RNA
CC can be used in a method for identifying compounds having antiviral
CC activity against HCV. When the pestivirus viral nucleotide sequence is
CC from bovine viral diarrhea virus (BVDV), the chimeric viral RNA can be
CC used in a vaccine against BVDV.
XX
SQ Sequence 3988 AA;
Query Match 30.1%; Score 81; DB 21; Length 3988;
Best Local Similarity 100.0%; Pred. No. 6.9e-68;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 103 MCSRCQKRRFFEMDRPKSARYCAECNRLHPAEEDGFWAESSMLGLKITFYALMDGKVY 162
DB 1537 MCSRCQKRRFFEMDRPKSARYCAECNRLHPAEEDGFWAESSMLGLKITFYALMDGKVY 1596
QY 163 DITEWAGCQRVGISPDTHRPV 183
DB 1597 DITEWAGCQRVGISPDTHRPV 1617
RESULT 13
ABB77045
ID ABB77045 standard; Protein; 204 AA.
XX
AC ABB77045;
XX
DT 08-OCT-2002 (first entry)
XX
DE Human protein sequence #2 from clone HNTPB82.
XX
KW Human; HNTPB82; secreted protein; immunosuppressive; food preservative;
KW antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;
KW vasotropic; cerebroprotective; nocotropic; neuroprotective; antibacterial;
KW virucide; fungicide; ophthalmological; vulnary; gene therapy; EUSA;
KW radioimmunoassay; enzyme linked immunosorbent assay; autoimmune disease;
KW hyperproliferative disorder; cardiovascular disorder; angiogenesis;
KW cerebrovascular disorder; nervous system disorder; ocular disorder;

KW wound healing; food additive.
XX Homo sapiens.
OS
XX
XX
FH Location/Qualifiers
FT Peptide 1..43
FT /label= Signal_peptide
FT Protein 44..204
FT /label= Mature_protein
XX WO200222638-A1.
XX
XX 21-MAR-2002.
XX
XX 17-JAN-2001; 2001WO-US01386.
XX
XX 12-SEP-2000; 2000US-232104P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;
XX Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;
XX Ni J;
XX WPI; 2002-258041/30.
XX N-PSDB; ABL55114.
XX
XX New nucleic acid molecules encoding 22 human secreted proteins for
XX diagnosing or treating e.g. autoimmune diseases, hyperproliferative
XX disorders, and cardiovascular disorders, and used as food additives or
XX preservatives -
XX
XX Disclosure; Page 507-508; 526pp; English.
XX
XX The sequence represents a protein sequence of the invention, encoded by
XX cDNA isolated from human clone ID HNPFB82. The invention relates to novel
XX isolated nucleic acid molecules encoding 22 human secreted proteins. The
XX proteins of the invention have immunosuppressive, antiarthritic,
XX antirheumatic, antiproliferative, cytostatic, cardiac, vasotropic,
XX cerebroprotective, neurotropic, neuroprotective, antibacterial, virucide,
XX fungicide, ophthalmological, and vulnery activity. The polynucleotides
XX may have a use in gene therapy. The polynucleotides and polypeptides
XX encoded by them are used to prevent, treat or ameliorate a medical
XX condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
XX chickens or sheep. The polynucleotides and polypeptides are also used in
XX diagnosing a pathological condition or susceptibility to a pathological
XX condition. The antibodies to the proteins can also be used in alleviating
XX symptoms associated with the disorders and in diagnostic immunoassays
XX e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA).
XX Disorders which are diagnosed or treated include autoimmune diseases,
XX hyperproliferative disorders, cardiovascular disorders, cerebrovascular
XX disorders, angiogenesis, nervous system disorders, infections caused by
XX bacteria, viruses and fungi and ocular disorders. The polypeptides can
XX also be used to aid wound healing and epithelial cell proliferation. The
XX polypeptides can also be used as a food additive or preservative.
XX
XX Sequence 204 AA;
SQ
Query Match 19.7%; Score 53; DB 23; Length 204;
Best Local Similarity 100.0%; Pred. No. 2.9e-42;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGVPEDELNPFVLGVETASDVELKKAYQLAVMVHPDKNHHPRAEAFKV 53
DB 130 MAGVPEDELNPFVLGVETASDVELKKAYQLAVMVHPDKNHHPRAEAFKV 182
RESULT 14
ABG64926
ID ABG64926 standard; Protein; 204 AA.
XX
XX AC ABG64926;
XX

DT 27-AUG-2002 (first entry)
XX
XX Human albumin fusion protein #1601.
XX
XX Albumin fusion protein; therapeutic protein X; human albumin; HA;
XX human serum albumin; HSA; cancer; reproductive disorder;
XX digestive disorder; immune disorder; endocrine disorder;
XX haematopoietic disorder; neural disorder; connective disorder;
XX cytostatic; antiinfertility; antiinflammatory; antiulcer;
XX immunomodulator; anti-HIV; antidiabetic; haemostatic; neurotropic;
XX neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
XX osteopathic; antiarthritic.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX WO200177137-A1.
XX
XX 18-OCT-2001.
XX
XX 12-APR-2001; 2001WO-US11988.
XX
XX 12-APR-2000; 2000US-229358P.
XX 25-APR-2000; 2000US-199384P.
XX 21-DEC-2000; 2000US-256931P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Haseltine WA;
XX WPI; 2002-010886/01.
XX
XX New fusion protein for treating disease e.g. diabetes comprises an
XX albumin fused to a therapeutic protein -
XX
XX Claim 1; Page 1605-1606; 2102pp; English.
XX
XX The present invention relates to albumin fusion proteins comprising a
XX therapeutic protein X and human albumin (HA, also known as human serum
XX albumin, HSA). The proteins are useful for treating a disease or
XX disorder that may be modulated by therapeutic protein X. The albumin
XX extends the shelf-life of protein X, and may increase its biological
XX in vitro/in vivo activity. The protein is useful for treating and
XX diagnosing disorders such as cancer, reproductive disorders, digestive
XX disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
XX (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
XX (e.g. diabetes), haematopoietic disorders, neural disorders
XX (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
XX encephalomyelitis, meningitis, schizophrenia), and connective disorders
XX (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin
XX fusion proteins of the invention.
XX
XX Sequence 204 AA;
SQ
Query Match 19.7%; Score 53; DB 23; Length 204;
Best Local Similarity 100.0%; Pred. No. 2.9e-42;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGVPEDELNPFVLGVETASDVELKKAYQLAVMVHPDKNHHPRAEAFKV 53
DB 130 MAGVPEDELNPFVLGVETASDVELKKAYQLAVMVHPDKNHHPRAEAFKV 182
RESULT 15
ABB69643
ID ABB69643 standard; Protein; 970 AA.
XX
XX AC ABB69643;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 35721.
XX

KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL13746.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

XX Disclosure; SEQ ID NO 35721; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 970 AA;

Query Match 5.2%; Score 14; DB 22; Length 970;

Best Local Similarity 100.0%; Pred. No. 0.00018;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 ALMDGKVYDITWA 168

Db 856 ALMDGKVYDITWA 869

Search completed: January 2, 2004, 16:45:28

Job time : 66 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 16:43:34 ; Search time 39 Seconds

(without alignments)
663.317 Million cell updates/sec

Title: US-10-049-742-11

Perfect score: 269

Sequence: 1 MAGVPEDELNPFHVLGEAT.....VPGGEAKPKRRKKVRRPQR 269

Scoring table:

OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

PIR 76:*

1: Piri:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	81	30.1	3988	1	GNWVBV	genome polyprotein
2	8	3.0	186	2	C95207	hypothetical prote
3	8	3.0	186	2	C98072	hypothetical prote
4	8	3.0	607	2	I37560	protein-tyrosine k
5	8	3.0	1474	1	MAHU	alpha-2-macroglobu
6	7	2.6	111	2	D53589	hypothetical prote
7	7	2.6	126	2	E75350	hypothetical prote
8	7	2.6	130	2	AG2468	50S ribosomal prot
9	7	2.6	149	2	S13459	hemoglobin - south
10	7	2.6	175	2	T23245	hypothetical prote
11	7	2.6	178	2	C45831	MHC class I histoc
12	7	2.6	189	2	S34632	dnaj protein homol
13	7	2.6	200	2	S67279	hypothetical prote
14	7	2.6	201	2	C70170	hypothetical prote
15	7	2.6	225	2	C84376	hypothetical prote
16	7	2.6	252	2	AD0267	TonB protein limpo
17	7	2.6	255	2	S30290	tonB protein - Yer
18	7	2.6	259	2	G70635	hypothetical prote
19	7	2.6	260	2	E69013	conserved hypother
20	7	2.6	299	2	S18813	homeotic protein H
21	7	2.6	309	2	C72589	hypothetical prote
22	7	2.6	323	2	S20099	transforming prote
23	7	2.6	323	2	A54772	homeotic protein M
24	7	2.6	324	2	T15283	hypothetical prote
25	7	2.6	324	2	D87725	protein M01D7.2 (i
26	7	2.6	335	2	T48161	heat shock protein
27	7	2.6	345	2	A53138	gas1 homolog - hum
28	7	2.6	354	2	S39406	homeotic protein o
29	7	2.6	362	2	A86831	X-Pro dipeptidase

30 7 2.6 379 2 S55900 DNaj-like protein
31 7 2.6 379 2 T41633 psi protein - f1ss
32 7 2.6 388 2 G75375 hypothetical prote
33 7 2.6 390 2 T10231 anther-specific pr
34 7 2.6 395 2 B95864 hypothetical prote
35 7 2.6 397 2 S34630 dnaj protein homol
36 7 2.6 401 2 A48423 engrailed homeodom
37 7 2.6 402 2 T21991 hypothetical prote
38 7 2.6 407 2 T39658 probable mitochond
39 7 2.6 410 2 C35971 probable sugar upt
40 7 2.6 412 2 I39554 phosphoglycerate k
41 7 2.6 413 2 I39551 phosphoglycerate k
42 7 2.6 418 2 T15142 hypothetical prote
43 7 2.6 427 2 B95046 trigger factor (im
44 7 2.6 427 2 B97917 trigger factor (im
45 7 2.6 432 2 F84747 probable SWI/SNF c

ALIGNMENTS

RESULT 1

GNWVBV
genome polyprotein - bovine viral diarrhea virus
C:Species: bovine viral diarrhea virus, BVDV
C:Date: 30-Sep-1990 #sequence_revision 30-Sep-1990 #text_change 19-Jan-2001
C:Accession: A29198; A61161
R:Collett, M.S.; Larson, R.; Gold, C.; Strick, D.; Anderson, D.K.; Purchio, A.F.
Virology 165, 191-199, 1988
A:Title: Molecular cloning and nucleotide sequence of the pestivirus bovine viral diarr
A:Reference number: A29198; MUID:88265858; PMID:2838957
A:Accession: A29198
A:Molecule type: genomic RNA
A:Residues: 1-3988 <COL>
A:Experimental source: isolate NADL
R:Ward, P.; Misra, V.
Am. J. Vet. Res. 52, 1231-1236, 1991
A:Title: Detection of bovine viral diarrhea virus, using degenerate oligonucleotide pri
A:Reference number: A61161; MUID:92027091; PMID:1656820
A:Accession: A61161
A:Molecule type: genomic RNA
A:Residues: 2054-2072 <WAR>
A:Experimental source: isolate V1352
A:Note: authors translated the codon ATA for residue 18 as Thr
C:Superfamily: pestivirus genome polyprotein
C:Keywords: ATP; Glycoprotein; nucleotide binding; P-loop; polyprotein
F:2-234/Product: viral proteinase p20 #status predicted <VPT>
F:548-1115/Product: major envelope glycoprotein gp55 #status predicted <EGP>
F:1905-1912/Region: nucleotide-binding motif A (P-loop)
F:1996-2001/Region: nucleotide-binding motif B
F:2000-2003/Region: DEXH motif
F:272,281,296,365,370,413,487,597,809,922,990,1357,1419,1451,1803,2224,2307,2584,2772,2:

Query Match 30.1%; Score 81; DB 1; Length 3988;
Best Local Similarity 100.0%; Pred. No. 9.7e-73;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 MCSRCQGHRRREMDREPKSARYCAECNRLHPAEGDFWAESSMLGLKITYPALMDGKVY 162
DB 1537 MCSRCQGHRRREMDREPKSARYCAECNRLHPAEGDFWAESSMLGLKITYPALMDGKVY 1596

QY 163 DITWAGCORVGISPDTHRPV 183

DB 1597 DITWAGCORVGISPDTHRPV 1617

RESULT 2

C95207
hypothetical protein SP1779 [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: C95207
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Hei

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzappple, nson, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Lofthus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A.
 A:Title: Complete Genome Sequence of a virulent isolate of *Streptococcus pneumoniae*.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: C95207
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-186 <KUR>
 A:Cross-references: GB:AE005672; PIDN:BAK75852.1; PID:gl4973275; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SPl779

Query Match 3.0%; Score 8; DB 2; Length 186;

Best Local Similarity 100.0%; Pred. No. 4.2;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 LQDDLKEA 98

Db 173 LQDDLKEA 180

RESULT 3

C98072 hypothetical protein spr1605 [imported] - *Streptococcus pneumoniae* (strain R6)

C:Species: *Streptococcus pneumoniae*

C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001

C:Accession: C98072

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; B

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M

y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A:Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6.

A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: C98072

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-186 <KUR>

A:Cross-references: GB:AE007317; PIDN:AAL00408.1; PID:gl5459273; GSPDB:GN00174

C:Genetics:

A:Gene: spr1605

Query Match 3.0%; Score 8; DB 2; Length 186;

Best Local Similarity 100.0%; Pred. No. 4.2;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 LQDDLKEA 98

Db 173 LQDDLKEA 180

RESULT 4

I37560 protein-tyrosine kinase (EC 2.7.1.112) ryk - human

C:Species: *Homo sapiens* (man)

C>Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 04-Feb-2000

C:Accession: I37560; A38269; S31579

R:Tamagnone, L.; Partanen, J.; Armstrong, E.; Lasota, J.; Ohgami, K.; Tazunoki, T.; LaFo

Onogene 8, 2009-2014, 1993

A:Title: The human ryk cDNA sequence predicts a protein containing two putative transmem

A:Reference number: I37560; MUID:93288416; PMID:8390040

A:Accession: I37560

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-607 <RES>

A:Cross-references: EMBL:X69970; NID:g32461; PIDN:CAA9591.1; PID:g32462

R:Partanen, J.; Maekelae, T.P.; Alitalo, R.; Lehtvaeslahti, H.; Alitalo, K.

Proc. Natl. Acad. Sci. U.S.A. 87, 8913-8917, 1990

A:Title: Putative tyrosine kinases expressed in K-562 human leukemia cells.

A:Reference number: A38268; MUID:91062389; PMID:2247464

A:Accession: A38269
 A:Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 467-523 <PAR>
 C:Genetics:
 A:Gene: GDB:RYK; D3S3195
 A:Cross-references: GDB:217730
 A:Map position: 3q22-3q22
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol
 C:Keywords: ATP; magnesium; phosphotransferase; tyrosine-specific protein kinase
 F:328-603/Domain: protein kinase homology <KIN>
 F:336-344/Region: protein kinase ATP-binding motif
 F:364,381,465/Active site: Lys, Glu, Asp #status predicted
 F:470,483/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 3.0%; Score 8; DB 2; Length 607;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 PAPGAAAA 243

Db 40 PAPGAAAA 47

RESULT 5

MAHU

alpha-2-macroglobulin precursor [invalidated] - human

N:Alternate names: alpha-2M

C:Species: *Homo sapiens* (man)

C>Date: 05-Apr-1983 #sequence_revision 30-Jun-1987 #text_change 08-Dec-2000

C:Accession: A94033; I39375; S09107; JN0262; A92486; S65634; A01256

R:Kan, C.C.; Solomon, E.; Beit, K.T.; Chain, A.C.; Hiorns, L.R.; Fey, G.

Proc. Natl. Acad. Sci. U.S.A. 82, 2282-2286, 1985

A:Title: Nucleotide sequence of cDNA encoding human alpha-2-macroglobulin and assignmen

A:Reference number: A94033; MUID:85190481; PMID:2581245

A:Accession: A94033

A:Molecule type: mRNA

A:Residues: 1-1474 <KAN>

A:Cross-references: GB:M11313; NID:gl77869; PIDN:AAA51551.1; PID:gl77870

A>Note: hydrolysis of the thiolester bond during amino acid sequencing of the mature pr

lymorphism

R:Bell, G.I.; Rall, L.B.; Sanchez-Pescador, R.; Merryweather, J.P.; Scott, J.; Eddy, R.

Somat. Cell Mol. Genet. 11, 285-289, 1985

A:Title: Human alpha 2-macroglobulin gene is located on chromosome 12.

A:Reference number: I39375; MUID:85219061; PMID:2408344

A:Accession: I39375

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 632-999, 'I', 1001-1147, 'D', 1149-1194, 'D', 1196-1474 <BEL>

A:Cross-references: GB:M36501; NID:gl77871; PIDN:AAA51552.1; PID:gl77872

R:Marynen, P.; Devriendt, K.; van den Berghe, H.; Cassiman, J.J.

FEBS Lett. 262, 349-352, 1990

A:Title: A genetic polymorphism in a functional domain of human pregnancy zone protein:

n.

A:Reference number: S09106; MUID:90242963; PMID:1692292

A:Accession: S09107

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 672-747 <MAR>

R:Mathijs, G.; Devriendt, K.; Cassiman, J.J.; Van Den Berghe, H.; Marynen, P.

Biochem. Biophys. Res. Commun. 184, 596-603, 1992

A:Title: Structure of the human alpha-2 macroglobulin gene and its promotor.

A:Reference number: JN0262; MUID:92246939; PMID:1374237

A:Accession: JN0262

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-29 <MAT>

R:Sostrup-Jensen, L.; Stepanik, T.M.; Kristensen, T.; Wierzbicki, D.M.; Jones, C.M.; Lo

J. Biol. Chem. 259, 8318-8327, 1984

A:Title: Primary structure of human alpha-2-macroglobulin. V. The complete structure.

A:Reference number: A92486; MUID:84239807; PMID:6203908

A:Accession: A92486

A:Molecule type: protein

A:Residues: 24-562, 'E', 564-974, 'Z', 976-999, 'I', 1001-1474 <SOT>
R:Sottrup-Jensen, L.; Stepanik, T.M.; Kristensen, T.; Wierzbicki, D.M.; Jones, C.M.; Lon
J. Biol. Chem. 260, 6500, 1985
A:Reference number: A92529
A:Contents: annotation; erratum
R:Virca, G.D.; Salvesen, G.S.; Travis, J.
Hoppe-Seyler's Z. Physiol. Chem. 364, 1297-1302, 1983
A:Title: Human neutrophil elastase and cathepsin G cleavage sites in the bait region of
A:Reference number: A91713; MUID:84030513; PMID:6195065
A:Contents: annotation; inhibitory site
R:Sottrup-Jensen, L.; Lonblad, P.B.; Stepanik, T.M.; Petersen, T.E.; Magnusson, S.; Jor
FEBS Lett. 127, 167-173, 1981
A:Title: Primary structure of the 'bait' region for proteinases in alpha-2-macroglobulin
A:Reference number: A91290; MUID:81212827; PMID:6165619
A:Contents: annotation; inhibitory site
R:Hall, P.K.; Nelles, L.P.; Travis, J.; Roberts, R.C.
Biochem. Biophys. Res. Commun. 100, 8-16, 1981
A:Title: Proteolytic cleavage sites on alpha-2-macroglobulin resulting in proteinase bin
A:Reference number: A90093; MUID:81255805; PMID:6167263
A:Contents: annotation; inhibitory site
R:Morlensen, S.B.; Sottrup-Jensen, L.; Hansen, H.F.; Petersen, T.E.; Magnusson, S.
FEBS Lett. 135, 295-300, 1981
A:Title: Primary and secondary cleavage sites in the bait region of alpha-2-macroglobulin
A:Reference number: A91299; MUID:82095610; PMID:6172288
A:Contents: annotation; inhibitory site
R:Dolner, K.; Jenner, L.B.; Jacobsen, L.; Andersen, G.R.; Koch, T.J.; Thirup, S.; Sottr
FEBS Lett. 372, 93-95, 1995
A:Title: Crystallisation and preliminary X-ray analysis of the receptor-binding domain o
A:Reference number: S66634; MUID:96032553; PMID:7556651
A:Accession: S66634
A:Molecule type: protein
A:Residues: 1337-1343 <DOL>
C:Comment: This inhibitor is able to inhibit all four classes (EC 3.4.21-3.4.24) of endo
s tertiary conformation, entrapping the proteinase. This results in hydrolysis of its th
e and appears to unmask a receptor-specific binding site
C:Comment: The entrapped enzyme remains active against low molecular weight substrates
r complex results in its rapid clearance from the circulation by receptor-mediated endoc
C:Comment: The wide specificity of this inhibitor is attributed to the primary sequence
ities of a variety of plasma proteinases, form primary (residues 704-809) and secondary
C:Genetics:
A:Gene: GDB:A2M
A:Cross-references: GDB:119639; OMIM:103950
A:Map position: 12p13.3-12p12.3
C:Complex: homotetramer; dimer of disulfide linked dimers
C:Superfamily: alpha-2-macroglobulin
C:Keywords: glycoprotein; plasma; proteinase inhibitor; thiolester bond
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-1474/Product: alpha-2-macroglobulin #status experimental <MA2>
F:693-694, 704-709, 719-723, 730-735/Region: inhibitory #status predicted
F:698-742/Region: bait region
F:48-86, 251-299, 269-287, 278-431, 595-771, 642-689, 821-849, 847-883, 921-1321, 1079-1127, 1352-
F:55, 70, 247, 396, 410, 869, 991, 1124/Binding site: carbohydrate (Asn) (covalent) #status exp
F:470, 563/Disulfide bonds: interchain #status predicted
F:972-975/Cross-link: thiolester (Cys-Gln) #status experimental

Query Match 3.0%; Score 8; DB 1; Length 1474;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 KRKEYEMK 73
|||||
DB 312 KRKEYEMK 319

RESULT 6
D53589
hypothetical protein beta-2 (B2 3'-region) - Carnobacterium piscicola (strain LV17B)
C:Species: Carnobacterium piscicola
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 04-Mar-2000
C:Accession: D53589
R:Quadr, L.E.N.; Sailer, M.; Roy, K.L.; Vederas, J.C.; Stiles, M.E.
J. Biol. Chem. 269, 12204-12211, 1994
A:Title: Chemical and genetic characterization of bacteriocins produced by Carnobacteriu

A:Reference number: A53589; MUID:94216339; PMID:8163526
A:Accession: D53589
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-111 <QUA>
A:Cross-references: GB:L29059; NID:9456409; PIDN:AAA72432.1; PID:9456411
C:Superfamily: Carnobacterium piscicola hypothetical protein beta-2 (B2 3'-region)

Query Match 2.6%; Score 7; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 NEFLSKL 91
|||||
DB 24 NEFLSKL 30

RESULT 7
E75350
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: E75350
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vanathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; M.
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: E75350
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-126 <WHI>
A:Cross-references: GB:AE002022; GB:AE000513; NID:G6459590; PIDN:AAF11378.1; PID:G64596
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1821
A:Map position: 1

Query Match 2.6%; Score 7; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 AQQPAPG 239
|||||
DB 73 AQQPAPG 79

RESULT 8
AG2468
50S ribosomal protein L12 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AG2468
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, ;
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An.
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2468
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-130 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA877002.1; PID:gl7134442; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: rpl12
C:Superfamily: Escherichia coli ribosomal protein L12

Query Match 2.6%; Score 7; DB 2; Length 130;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 APOGAAA 243
Db 47 APOGAAA 53

RESULT 9
hemoglobin - southern lamprey
S13459
C:Species: Mordacia mordax (southern lamprey)
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 04-Mar-2000
C:Accession: S13459
R:Hombrados, I.; Vidal, Y.; Rodewald, K.; Braunitzer, G.; Neuzil, E.
Biol. Chem. Hoppe-Seyler 372, 49-56, 1991
A:Title: The primary structure of the hemoglobins of a southern hemisphere lamprey (Mordacia mordax)
A:Reference number: S13458; MUID:91248417; PMID:2039605
A:Accession: S13459
A:Molecule type: protein
A:Residues: 1-149 <HOM>
C:Function:
C:Superfamily: globin; globin homology
C:Keywords: chromoprotein; erythrocyte; heme; iron; metalloprotein; oxygen carrier
F;11-149/Domain: globin homology <GLB>
F;73/Binding site: oxygen (His) (distal axial ligand) #status predicted
F;105/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 2.6%; Score 7; DB 2; Length 149;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 RAAWDIV 61
Db 20 RAAWDIV 26

RESULT 10
T23245
hypothetical protein K02E11.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T23245
R:McMurray, A.
submitted to the EMBL Data Library, July 1996
A:Reference number: Z19715
A:Accession: T23245
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-175 <WIL>
A:Cross-references: EMBL:Z77665; PIDN:CAB01221.1; GSPDB:GN00023; CESP:K02E11.6
A:Experimental source: clone K02E11
C:Genetics:
A:Gene: CESP:K02E11.6
A:Map position: 5
A:Introns: 30/3; 79/3; 128/2

Query Match 2.6%; Score 7; DB 2; Length 175;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 SKPNSTV 250
Db 29 SKPNSTV 35

RESULT 11
C45831
MHC class I histocompatibility antigen alpha chain (clone SC116), extended form - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 23-Jul-1999
C:Accession: C45831
R:Grosberger, D.; Hein, W.; Marcuz, A.
Immunogenetics 32, 77-87, 1990

A:Title: Class I major histocompatibility complex cDNA clones from sheep thymus: alternative splicing
A:Reference number: A45831; MUID:90375196; PMID:2397935
A:Accession: C45831
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-178 <GRO>
A:Cross-references: GB:M34674; NID:gl65991; PIDN:AAA31566.1; PID:gl65992
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;14-79/Domain: immunoglobulin homology <IMM>

Query Match 2.6%; Score 7; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 TVPKGEA 255
Db 151 TVPKGEA 157

RESULT 12
S34632
dnaJ protein homolog - human
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Dec-1997
C:Accession: S34632
R:Chellaiyah, A.; Davis, A.R.; Mohanakumar, T.
submitted to the EMBL Data Library, December 1992
A:Description: Cloning of a unique human homologue of the Escherichia coli DnaJ heat shock protein
A:Reference number: S34632
A:Accession: S34632
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-189 <CHE>
A:Cross-references: EMBL:L08069
C:Superfamily: dnaJ amino-terminal homology
F;6-68/Domain: dnaJ amino-terminal homology <DNJ>

Query Match 2.6%; Score 7; DB 2; Length 189;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 ELKKAYR 31
Db 21 ELKKAYR 27

RESULT 13
S67279
hypothetical protein YOR367w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein O6667
C:Species: Saccharomyces cerevisiae
C>Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 20-Jun-2000
C:Accession: S67279
R:Delius, H.; Hebling, U.; Hofmann, B.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67261
A:Accession: S67279
A:Molecule type: DNA
A:Residues: 1-200
A:Cross-references: EMBL:Z75275; NID:gl420793; PID:gl420794; MIPS:YOR367w
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:SCP1
A:Cross-references: SGD:S0005994; MIPS:YOR367w
A:Map position: 15R
C:Superfamily: smooth muscle protein SM22 homology <SMH>
F;10-200/Domain: smooth muscle protein SM22 homology <SMH>

Query Match 2.6%; Score 7; DB 2; Length 200;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GVPEDEL 9

```
Db      104 GVPEDL 110
|||||
RESULT 14
C70170
hypothetical protein BB0564 - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: C70170
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: C70170
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-201 <KLE>
A:Cross-references: GB:AE001158; GB:AE000783; NID:G2688483; PIDN:AAC66939.1; PID:G268849
A:Experimental source: strain B31
Query Match      2.6%; Score 7; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 45;
Matches      7; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
QY      223 GQMPNGN 229
Db      176 GQMPNGN 182
|||||
RESULT 15
C84376
hypothetical protein vng2255c [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: C84376
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabc
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: C84376
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-225 <STO>
A:Cross-references: GB:AE004437; NID:g10581668; PIDN:AAG20375.1; GSPDB:GN00138
C:Genetics:
A:Gene: VNG2255C
C:Superfamily: Methanobacterium thermoautotrophicum phosphatidylserine decarboxylase
Query Match      2.6%; Score 7; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 50;
Matches      7; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
QY      185 HISFGSR 191
Db      188 HISFGSR 194
|||||
Search completed: January 2, 2004, 16:47:25
Job time : 40 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 16:42:34 ; Search time 17 Seconds
(without alignments)
744.128 Million cell updates/sec

Title: US-10-049-742-11
Perfect score: 269
Sequence: 1 MAGVPEDELNPFHVLGVEAT.....VPGKEAKPKRKKVRRPFQR 269

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB ID	Description
	Score	Match	Length		
1	81	30.1	3988	1	POLG BVDVN
2	8	3.0	309	1	SCA1_TRIVA
3	8	3.0	309	1	SCA2_TRIVA
4	8	3.0	309	1	SCA3_TRIVA
5	8	3.0	370	1	DNAJ_EYRH
6	8	3.0	522	1	SNX1_RAT
7	8	3.0	1474	1	A2MG_HUMAN
8	7	2.6	100	1	SUL1_SULTO
9	7	2.6	111	1	CB21_CARPI
10	7	2.6	129	1	RL7_ANASP
11	7	2.6	149	1	GLB2_MORMR
12	7	2.6	255	1	TONB_YEREN
13	7	2.6	293	1	MSX1_MOUSE
14	7	2.6	323	1	JUND_CHICK
15	7	2.6	345	1	GAS1_HUMAN
16	7	2.6	354	1	OTX1_HUMAN
17	7	2.6	369	1	DNAJ_NITEU
18	7	2.6	370	1	MYOM_APLCA
19	7	2.6	379	1	PSI_SCHPO
20	7	2.6	383	1	DNAJ_FORGI
21	7	2.6	386	1	HXAD_MOUSE
22	7	2.6	397	1	DJA1_HUMAN
23	7	2.6	397	1	DJA1_MOUSE
24	7	2.6	397	1	PGK_CLOPE
25	7	2.6	401	1	HME1_MOUSE
26	7	2.6	412	1	DJA2_HUMAN
27	7	2.6	412	1	DJA2_MOUSE
28	7	2.6	412	1	DJA2_RAT
29	7	2.6	412	1	PGKP_ALCEU
30	7	2.6	413	1	PGKC_ALCEU
31	7	2.6	427	1	TTC_STRPN
32	7	2.6	431	1	MYCS_MOUSE
33	7	2.6	450	1	G6PI_LISIN

34	7	2.6	450	1	G6PI_LISMO
35	7	2.6	481	1	ATPB_MESVI
36	7	2.6	514	1	C11A_DASAM
37	7	2.6	522	1	LEU1_DEIRA
38	7	2.6	554	1	PYRG_FALSO
39	7	2.6	590	1	YNW7_YEAST
40	7	2.6	599	1	LEPA_UREPA
41	7	2.6	609	1	OPA_DROME
42	7	2.6	686	1	VGLH_PVKA
43	7	2.6	686	1	VGLH_PVYN3
44	7	2.6	686	1	VGLH_PVRI
45	7	2.6	745	1	NLDL_RAT

ALIGNMENTS

RESULT 1
ID POLG BVDVN STANDARD; PRT; 3988 AA.
AC P19711,
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Genome polyprotein.
OS Bovine viral diarrhea virus (isolate NADL) (BVDV) (Mucosal disease virus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Pestivirus.
OX NCBI_TaxID=11100;
RN [1]
RP SEQUENCE FROM N.A. PubMed=2838957;
RX MEDLINE=88265858; PubMed=2838957;
RA Collett M.S., Larson R., Gold C., Strick D., Anderson D.K.,
RA Purchio A.F.;
RT "Molecular cloning and nucleotide sequence of the pestivirus bovine
RT viral diarrhea virus."
RL Virology 165:191-199(1998).
RN [2]
RP GENOMIC ORGANIZATION.
RX MEDLINE=88265859; PubMed=2838958;
RA Collett M.S., Larson R., Belzer S.K., Retzel E.;
RT "Proteins encoded by bovine viral diarrhea virus: the genomic
RT organization of a pestivirus."
RL Virology 165:200-208(1998).
CC -!- FUNCTION: PESTIVIRUS P80 (P125) MAY BE A BIFUNCTIONAL PROTEIN
CC WITH HELICASE AND PROTEASE ACTIVITY.
CC -!- PTM: GP116 GIVES RISE TO GP62 AND GP53; GP62 IN TURN YIELDS GP48
CC AND GP25.
CC -!- SIMILARITY: TO THE HOG CHOLERA VIRUS GENOME POLYPROTEIN.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S31.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M31182; AAA42854.1; -
DR PIR; A29198; GNMVBU.
DR HSSP; P27958; 1A1V.
DR MEROPS; C53.001; -
DR MEROPS; S31.001; -
DR InterPro; IPR000280; CDvir_endptsep80.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase.C.
DR InterPro; IPR007095; RNA_pol_PS-PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR InterPro; IPR001568; RNase T2.
DR Pfam; PF00271; helicase_C; 1.

PFam; PF00998; Viral RGRP; 1.
DR PRINTS; PR00729; CDVENDOPTASE.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00531; RNASE T2 2; UNKNOWN 1.
KW Polyprotein; Glycoprotein; Helicase; Serine protease; Hydrolase.
FT CHAIN 1 ?270
FT CHAIN ?271 ?1063
FT CHAIN ? ?
FT CHAIN ? ? 3988
FT CARBOHYD 272 272
FT CARBOHYD 281 281
FT CARBOHYD 296 296
FT CARBOHYD 335 335
FT CARBOHYD 365 365
FT CARBOHYD 370 370
FT CARBOHYD 413 413
FT CARBOHYD 457 457
FT CARBOHYD 597 597
FT CARBOHYD 809 809
FT CARBOHYD 878 878
FT CARBOHYD 922 922
FT CARBOHYD 990 990
FT CARBOHYD 1357 1357
FT CARBOHYD 1419 1419
FT CARBOHYD 1451 1451
FT CARBOHYD 1803 1803
FT CARBOHYD 2224 2224
FT CARBOHYD 2307 2307
FT CARBOHYD 2584 2584
FT CARBOHYD 2772 2772
FT CARBOHYD 2981 2981
FT CARBOHYD 3778 3778
FT CARBOHYD 3867 3867
FT CARBOHYD 3883 3883
FT CARBOHYD 3988 3988
SQ SEQUENCE 3988 AA; 449154 MW; 4474212F338661B8 CRC64;

Query Match 30.1%; Score 81; DB 1; Length 3988;
Best Local Similarity 100.0%; Pred. No. 7.6e-73;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 MCSRCQKRRRMDREPKSARYCAECNRLHPAEGDFWAESSMLGLKITYFALMDGKY 162
DB 1537 MCSRCQKRRRMDREPKSARYCAECNRLHPAEGDFWAESSMLGLKITYFALMDGKY 1596

QY 163 DITEWAGCORVGISPDTHRPV 183
DB 1597 DITEWAGCORVGISPDTHRPV 1617

RESULT 2
SCAL_TRIVA STANDARD; PRT; 309 AA.
AC P53399;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Succinyl-CoA ligase [GDP-forming] alpha-chain 1 precursor (EC 6.2.1.4)
DE (Succinyl-CoA synthetase, alpha chain 1).
GN ALPHA-SCS1.
OS Trichomonas vaginalis.
OC Eukaryota; Parabasalidea; Trichomonadida; Trichomonadidae;
OC Trichomonadinae; Trichomonas.
OX NCBI_TaxID=5722;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 30001 / NIH-C1;
RX MEDLINE=95107353; PubMed=7808480;
RA Lahti C.J., Bradley P.J., Johnson P.J.;
RT "Molecular characterization of the alpha-subunit of Trichomonas vaginalis hydrogenosomal succinyl CoA synthetase.";
RL Mol. Biochem. Parasitol. 66:309-318(1994).
CC -I- CATALYTIC ACTIVITY: GTP + succinate + CoA = GDP + succinyl-CoA + phosphate.

phosphate.
-I- PATHWAY: Tricarboxylic acid cycle.
-I- SUBUNIT: Tetramer of two alpha and two beta chains (By similarity).
-I- SUBCELLULAR LOCATION: Hydrogenosomal.
-I- SIMILARITY: Belongs to the succinate/malate CoA ligase alpha subunit family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; L31929; AAC41558.1; -
HSSP; P07459; 1CQJ.
DR InterPro; IPR003781; CoA binding.
DR InterPro; IPR005810; CoA_lig_alpha.
DR InterPro; IPR005811; CoA_ligase.
DR Pfam; PF02629; CoA_binding; 1.
DR Pfam; PF00549; ligase-CoA; 1.
DR TIGRFAMs; TIGR01019; sucCoAalpha; 1.
DR PROSITE; PS01216; SUCCINYL_COA_Lig_1; 1.
DR PROSITE; PS00399; SUCCINYL_COA_Lig_2; 1.
KW Ligase; Glycolysis; Tricarboxylic acid cycle; GTP-binding;
KW Phosphorylation; Hydrogenosome; Multigene family.
FT PROPEP 1 ?
FT CHAIN ? 309 SUCCINYL-CoA LIGASE [GDP-FORMING] ALPHA-CHAIN 1.
FT ACT_SITE 262 262 PHOSPHOHISTIDINE INTERMEDIATE (BY SIMILARITY);
FT SEQUENCE 309 AA; 32408 MW; 47FB103451F1E5E3 CRC64;

Query Match 3.0%; Score 8; DB 1; Length 309;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 PAPGAAA 243
DB 85 PAPGAAA 92

RESULT 3
SCA2_TRIVA STANDARD; PRT; 309 AA.
ID SCA2 TRIVA
AC P53400;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Succinyl-CoA ligase [GDP-forming] alpha-chain 2 precursor (EC 6.2.1.4)
DE (Succinyl-CoA synthetase, alpha chain 2).
GN ALPHA-SCS2.
OS Trichomonas vaginalis.
OC Eukaryota; Parabasalidea; Trichomonadida; Trichomonadidae;
OC Trichomonadinae; Trichomonas.
OX NCBI_TaxID=5722;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 30001 / NIH-C1;
RX MEDLINE=95107353; PubMed=7808480;
RA Lahti C.J., Bradley P.J., Johnson P.J.;
RT "Molecular characterization of the alpha-subunit of Trichomonas vaginalis hydrogenosomal succinyl CoA synthetase.";
RL Mol. Biochem. Parasitol. 66:309-318(1994).
CC -I- CATALYTIC ACTIVITY: GTP + succinate + CoA = GDP + succinyl-CoA + phosphate.
-I- PATHWAY: Tricarboxylic acid cycle.
-I- SUBUNIT: Tetramer of two alpha and two beta chains (By similarity).
-I- SUBCELLULAR LOCATION: Hydrogenosomal.
-I- SIMILARITY: Belongs to the succinate/malate CoA ligase alpha

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; L31931; AAC41560.1; --
HSSP; P07459; 1SCU.
InterPro; IPR003781; CoA binding.
InterPro; IPR005810; CoA_lig_alpha.
InterPro; IPR005811; CoA_ligase.
Pfam; PF02629; CoA binding; 1.
Pfam; PF00549; ligase-CoA; 1.
TIGRFAMs; TIGR01019; succoalalpha; 1.
PROSITE; PS01216; SUCCINYL_COA_LIG_1; 1.
PROSITE; PS00399; SUCCINYL_COA_LIG_2; 1.
Ligase; Glycolysis; Tricarboxylic acid cycle; GTP-binding;
Phosphorylation; Hydrogenosome; Multigene family.
PROPEP 1 ?
CHAIN 1 ? 309 SUCCINYL-COA LIGASE [GDP-FORMING] ALPHA-
ACT_SITE 262 262 PHOSPHOHISTIDINE INTERMEDIATE (BY
SEQUENCE 309 AA; 32513 MW; 966277FA7AF6768 CRC64;
Query Match 3.0%; Score 8; DB 1; Length 309;
Best Local Similarity 100.0%; Pred.No.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 PAPGAAAA 243
Db 85 PAPGAAAA 92

RESULT 5
DNAJ_ERYRH
ID DNAJ_ERYRH STANDARD; PRT; 370 AA.
AC Q05646;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chaperone protein dnaJ.
GN DNAJ.
OS Erysielothrix rhusiopathiae.
OC Bacteria; Firmicutes; Mollicutes; Anaeroplasmatales;
OC Erysielothrix rhusiopathiae; Erysielothrix.
OX NCBI_TaxID=1648;
RN [1] _
RP SEQUENCE FROM N.A.
RC STRAIN=EL-6P;
RX MEDLINE=93366167; PubMed=8359682;
RA Rockabrand D., Partridge J., Krska J., Blum P.;
RT "Nucleotide sequence analysis and heterologous expression of the
Erysielothrix rhusiopathiae dnaJ gene";
RL FEMS Microbiol. Lett. 111:79-85(1993).
CC -!- FUNCTION: ACTS AS A CO-CHAPERONE. STIMULATES, JOINTLY WITH GRPE,
THE ATPASE ACTIVITY OF DNAK (BY SIMILARITY).
CC -!- COFACTOR: BINDS TWO ZINC IONS PER MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- INDUCTION: By heat shock.
CC -!- SIMILARITY: BELONGS TO THE DNAJ FAMILY.
CC -!- SIMILARITY: Contains 1 J domain.
CC -!- SIMILARITY: Contains 1 CR domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; L08110; AAA71922.1; --
HSSP; P08622; 1BQZ.
InterPro; IPR002939; DnaJ_C.

subunit family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; L31930; AAC41559.1; --
HSSP; P07459; 1SCU.
InterPro; IPR003781; CoA binding.
InterPro; IPR005810; CoA_lig_alpha.
InterPro; IPR005811; CoA_ligase.
Pfam; PF02629; CoA binding; 1.
Pfam; PF00549; ligase-CoA; 1.
TIGRFAMs; TIGR01019; succoalalpha; 1.
PROSITE; PS01216; SUCCINYL_COA_LIG_1; 1.
PROSITE; PS00399; SUCCINYL_COA_LIG_2; 1.
Ligase; Glycolysis; Tricarboxylic acid cycle; GTP-binding;
Phosphorylation; Hydrogenosome; Multigene family.
PROPEP 1 ?
CHAIN 1 ? 309 SUCCINYL-COA LIGASE [GDP-FORMING] ALPHA-
ACT_SITE 262 262 PHOSPHOHISTIDINE INTERMEDIATE (BY
SEQUENCE 309 AA; 32653 MW; EAE674CEFCF3A90 CRC64;
Query Match 3.0%; Score 8; DB 1; Length 309;
Best Local Similarity 100.0%; Pred.No.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 PAPGAAAA 243
Db 85 PAPGAAAA 92

RESULT 4
SCA3_TRIVA
ID SCA3_TRIVA STANDARD; PRT; 309 AA.
AC P53401;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Succinyl-CoA ligase [GDP-forming] alpha-chain 3 precursor (EC 6.2.1.4)
DE (Succinyl-CoA synthetase, alpha chain 3).
GN ALPHA-SCS3.
OS Trichomonas vaginalis.
OC Eukaryota; Parabasalidea; Trichomonadida; Trichomonadidae;
OC Trichomonadinae; Trichomonas.
OX NCBI_TaxID=5722;
RN [1] _
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 30001 / NIH-C1;
RX MEDLINE=95107353; PubMed=7808480;
RA Lahiri C.J., Bradley P.J., Johnson P.J.;
RT "Molecular characterization of the alpha-subunit of Trichomonas
vaginalis hydrogenosomal succinyl CoA synthetase";
RL Mol. Biochem. Parasitol. 66:309-318(1994).
CC -!- CATALYTIC ACTIVITY: GTP + succinate + CoA = GDP + succinyl-CoA +
phosphate.
CC -!- PATHWAY: Tricarboxylic acid cycle.
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains
(BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Hydrogenosomal.
CC -!- SIMILARITY: Belongs to the succinate/malate CoA ligase alpha
subunit family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way

DR InterPro: IPR001305; DnaJ_CXXCXGKG.
DR InterPro: IPR001623; DnaJ_N.
DR InterPro: IPR003095; Hsp_DnaJ.
DR Pfam: PF00226; DnaJ_1.
DR Pfam: PF01556; DnaJ_C; 1.
DR Pfam: PF00684; DnaJ_CXXCXGKG; 1.
DR PRINTS: PR00625; DNAJPROTEIN.
DR SMART: SMO0271; DnaJ; 1.
DR PROSITE: PS00636; DnaJ_1; 1.
DR PROSITE: PS00076; DnaJ_2; 1.
DR PROSITE: PS00637; DnaJ_CXXCXGKG; 1.
KW Chaperone; DNA replication; Heat shock; Repeat; Zinc; Metal-binding.
FT DOMAIN 4 72
FT DOMAIN 77 106
FT REPEAT 147 154
FT REPEAT 164 171
FT REPEAT 190 197
FT REPEAT 204 211
FT REPEAT 147 147
FT METAL 150 150
FT METAL 164 164
FT METAL 167 167
FT METAL 190 190
FT METAL 193 193
FT METAL 204 204
FT METAL 207 207
SQ SEQUENCE 370 AA; 40331 MW; 9A2B0F0E4546767D CRC64;

Query Match 3.0%; Score 8; DB 1; Length 370;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 KKAYROLA 34
DB 23 KKAYROLA 30

RESULT 6
SNX1 RAT
ID -SNX1 RAT STANDARD; PRT; 522 AA.
AC Q99N27;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sorting nexin 1.
GN SNX1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11110793;
RA Chin L.S., Raynor M.C., Wei X., Chen H.Q., Li L.;
RT "Hrs interacts with sorting Nexin 1 and regulates degradation of
epidermal growth factor receptor.";
RL J. Biol. Chem. 276:7069-7078(2001).
CC -!- FUNCTION: MAY BE INVOLVED IN SEVERAL STAGES OF INTRACELLULAR
TRAFFICKING. PLAYS A ROLE IN TARGETING LIGAND-ACTIVATED EGFR TO
THE LYOSOMES FOR DEGRADATION AFTER ENDOCYTOSIS FROM THE CELL
SURFACE AND RELEASE FROM THE GOLGI.
CC -!- SIMILARITY: BELONGS TO THE SORTING NEXIN FAMILY.
CC -!- SIMILARITY: Contains 1 phox homology (PX) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; AF218916; AAG59616.1; -;
DR

DR InterPro: IPR001683; PX.
DR InterPro: IPR005329; Sorting_nexin_N.
DR Pfam: PF00787; PX; 1.
DR Pfam: PF03700; Sorting_nexin; 1.
DR SMART: SMO0312; PX; 1.
DR PROSITE: PS00195; PX; 1.
KW Transport; Protein transport; Golgi stack.
FT DOMAIN 143 272
FT DOMAIN 522 AA; 59044 MW; F0F05664087E4D24 CRC64;

Query Match 3.0%; Score 8; DB 1; Length 522;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 239 GAAAKSKP 246
DB 49 GAAAKSKP 56

RESULT 7
A2MG HUMAN
ID -A2MG HUMAN STANDARD; PRT; 1474 AA.
AC P01023; Q13677;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Alpha-2-macroglobulin precursor (Alpha-2-M).
GN A2M.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85190481; PubMed=2581245;
RA Kan C.-C., Solomon E., Belt K.T., Chain A.C., Hioris L.R., Fey G.H.;
RT "Nucleotide sequence of cDNA encoding human alpha 2-macroglobulin and
assignment of the chromosomal locus.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2282-2286(1985).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT HIS-704.
RX MEDLINE=92246939; PubMed=1374237;
RA Matthijs G., Devriendt K., Cassiman J.-J., van den Berghe H.,
RA Marynen P.;
RT "Structure of the human alpha-2 macroglobulin gene and its promotor.";
RL Biochem. Biophys. Res. Commun. 184:596-603(1992).
RN [3]
RP SEQUENCE OF 24-1474, AND CHARACTERIZATION.
RX MEDLINE=84239807; PubMed=6203908;
RA Sottrup-Jensen L., Stepanik T.M., Kristensen T., Wierzbicki D.M.,
RA Jones C.M., Loenblad P.B., Magnusson S., Petersen T.E.;
RT "Primary structure of human alpha 2-macroglobulin. V. The complete
structure.";
RL J. Biol. Chem. 259:8318-8327(1984).
RN [4]
RP ERRATUM.
RA Sottrup-Jensen L., Stepanik T.M., Kristensen T., Wierzbicki D.M.,
RA Jones C.M., Loenblad P.B., Magnusson S., Petersen T.E.;
RL J. Biol. Chem. 260:6500-6500(1985).
RN [5]
RP SEQUENCE OF 832-1474 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=85219061; PubMed=2408344;
RA Bell G.I., Rall L.B., Sanchez-Pescador R., Merryweather J.P.,
RA Scott J., Eddy R.L., Shows T.B.;
RT "Human alpha 2-macroglobulin gene is located on chromosome 12.";
RL Sonat. Cell Mol. Genet. 11:285-289(1985).
RN [6]
RP SEQUENCE OF 672-747
RX MEDLINE=90242963; PubMed=1692292;
RA Marynen P., Devriendt K., van den Berghe H., Cassiman J.J.;
RT "A genetic polymorphism in a functional domain of human pregnancy
zone protein: the bait region. Genomic structure of the bait domains
RT

[illegible]

INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).

EMBL; M11913; AAS1551.1; --
EMBL; Z11711; CAAT7774.1; --
EMBL; M36501; AAS1552.1; --
EMBL; X68728; CAA48670.1; --
EMBL; X68729; CAA48670.1; JOINED.
PIR; A94033; MAHU.
PDB; 1BV8; 30-SEP-98.
DR SWISS-2DPAGE; P01023; HUMAN.
DR Genew; HGNC:7; A2M.
DR MIM; 103950; --
GO; GO:0016975; F:alpha-2 macroglobulin; NAS.
InterPro; IPR002890; A2M_N.
DR InterPro; IPR001599; MacroglblnA2.
DR Pfam; PF00207; A2M; 1.
DR Pfam; PF01835; A2M.N; 1.
DR PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
KW Serine protease inhibitor; Glycoprotein; Plasma; Bait region; Signal;
3D-structure; Polymorphism; Thioester bond.
FT SIGNAL 1 23
FT CHAIN 24 1474
FT SITE 693 694

FT DOMAIN 690 728
FT SITE 704 709
FT SITE 719 723
FT SITE 730 735
FT DISULFID 48 86
FT DISULFID 251 299
FT DISULFID 269 287
FT DISULFID 278 431
FT DISULFID 470 470
FT DISULFID 563 563
FT DISULFID 595 771
FT DISULFID 642 689
FT DISULFID 821 849
FT DISULFID 847 883
FT DISULFID 921 1321
FT DISULFID 1079 1127
FT DISULFID 1352 1467
FT CROSSLINK 972 975
FT CARBOHYD 55 55
FT CARBOHYD 70 70
FT CARBOHYD 247 247
FT CARBOHYD 396 396
FT CARBOHYD 410 410
FT CARBOHYD 869 869
FT CARBOHYD 991 991
FT CARBOHYD 1424 1424
FT VARIANT 704 704

FT VARIANTS TYR-972 AND ILE-1000.
MEDLINE=92128897; PubMed=1370808;
Poller W., Faber J.-P., Klobbeck G., Olek K.;
"Cloning of the human alpha 2-macroglobulin gene and detection of mutations in two functional domains: the bait region and the thiolester site."
Hum. Genet. 88:313-319(1992).
CC -!- FUNCTION: IS ABLE TO INHIBIT ALL FOUR CLASSES OF PROTEINASES BY A UNIQUE "TRAPPING" MECHANISM. THIS PROTEIN HAS A PEPTIDE STRETCH, CALLED THE 'BAIT REGION' WHICH CONTAINS SPECIFIC CLEAVAGE SITES FOR DIFFERENT PROTEINASES. WHEN A PROTEINASE CLEAVES THE BAITS IN THE 'BAIT REGION', A CONFORMATIONAL CHANGE IS INDUCED IN THE PROTEIN WHICH TRAPS THE ENZYME. THE ENTRAPPED ENZYME REMAINS ACTIVE AGAINST LOW MOLECULAR WEIGHT SUBSTRATES (ACTIVITY AGAINST HIGH MOLECULAR WEIGHT SUBSTRATES IS GREATLY REDUCED). FOLLOWING CLEAVAGE IN THE BAITS, A THIOLESTER BOND IS HYDROLYZED AND MEDIATES THE COVALENT BINDING OF THE PROTEIN TO THE PROTEINASE.
SUBUNIT: HOMOTETRAMER, WHICH CONSISTS OF TWO PAIRS OF DISULPHIDE-LINKED CHAINS.
TISSUE SPECIFICITY: Plasma.
MISCELLANEOUS: CONTRARY TO THE RAT PROTEIN, WHICH IS AN ACUTE PHASE PROTEIN, THIS PROTEIN IS ALWAYS PRESENT AT HIGH LEVELS IN CIRCULATION.
SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,
of human pregnancy zone protein and alpha 2 macroglobulin." ;
FEBS Lett. 262:349-352(1990).
[7]
INHIBITORY SITE.
MEDLINE=84030513; PubMed=6195065;
Virca G.D., Salvanesen G.S., Travis J.;
"Human neutrophil elastase and cathepsin G cleavage sites in the bait region of alpha 2-macroglobulin. Proposed structural limits of the bait region."
Hoppe-Seyler's Z. Physiol. Chem. 364:1297-1302(1983).
[8]
INHIBITORY SITE.
MEDLINE=81212827; PubMed=6165619;
Sottrup-Jensen L., Loebblad P.B., Stepanik T.M., Petersen T.E., Magnusson S., Joernvall H.;
"Primary structure of the 'bait' region for proteinases in alpha 2-macroglobulin. Nature of the complex."
FEBS Lett. 127:167-173(1981).
[9]
INHIBITORY SITE.
MEDLINE=81255805; PubMed=6167263;
Hall P.K., Nelles L.P., Travis J., Roberts R.C.;
"Proteolytic cleavage sites on alpha 2-macroglobulin resulting in proteinase binding are different for trypsin and Staphylococcus aureus V-8 proteinase."
Biochem. Biophys. Res. Commun. 100:8-16(1981).
[10]
INHIBITORY SITE.
MEDLINE=82095610; PubMed=6172288;
Mortenson S.B., Sottrup-Jensen L., Hansen H.F., Petersen T.E., Magnusson S.;
"Primary and secondary cleavage sites in the bait region of alpha 2-macroglobulin."
FEBS Lett. 135:295-300(1981).
[11]
STRUCTURE BY NMR OF 1337-1474.
MEDLINE=99081553; PubMed=9865955;
Huang W., Dolmer K., Liaw X., Gettings P.G.W.;
"Localization of basic residues required for receptor binding to the single alpha-helix of the receptor binding domain of human alpha2-macroglobulin."
Protein Sci. 7:2602-2612(1998).
[12]
VARIANT ILE-1000.
MEDLINE=91187639; PubMed=1707161;
Poller W., Faber J.-P., Olek K.;
"Sequence polymorphism in the human alpha2-macroglobulin (A2M) gene."
Nucleic Acids Res. 19:198-198(1991).
[13]
VARIANTS TYR-972 AND ILE-1000.
MEDLINE=92128897; PubMed=1370808;
Poller W., Faber J.-P., Klobbeck G., Olek K.;
"Cloning of the human alpha 2-macroglobulin gene and detection of mutations in two functional domains: the bait region and the thiolester site."
Hum. Genet. 88:313-319(1992).
CC -!- FUNCTION: IS ABLE TO INHIBIT ALL FOUR CLASSES OF PROTEINASES BY A UNIQUE "TRAPPING" MECHANISM. THIS PROTEIN HAS A PEPTIDE STRETCH, CALLED THE 'BAIT REGION' WHICH CONTAINS SPECIFIC CLEAVAGE SITES FOR DIFFERENT PROTEINASES. WHEN A PROTEINASE CLEAVES THE BAITS IN THE 'BAIT REGION', A CONFORMATIONAL CHANGE IS INDUCED IN THE PROTEIN WHICH TRAPS THE ENZYME. THE ENTRAPPED ENZYME REMAINS ACTIVE AGAINST LOW MOLECULAR WEIGHT SUBSTRATES (ACTIVITY AGAINST HIGH MOLECULAR WEIGHT SUBSTRATES IS GREATLY REDUCED). FOLLOWING CLEAVAGE IN THE BAITS, A THIOLESTER BOND IS HYDROLYZED AND MEDIATES THE COVALENT BINDING OF THE PROTEIN TO THE PROTEINASE.
SUBUNIT: HOMOTETRAMER, WHICH CONSISTS OF TWO PAIRS OF DISULPHIDE-LINKED CHAINS.
TISSUE SPECIFICITY: Plasma.
MISCELLANEOUS: CONTRARY TO THE RAT PROTEIN, WHICH IS AN ACUTE PHASE PROTEIN, THIS PROTEIN IS ALWAYS PRESENT AT HIGH LEVELS IN CIRCULATION.
SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,

FT STRAND 1379 1384
FT STRAND 1390 1391

Query Match 3.0%; Score 8; DB 1; Length 1474;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 66 KRKEYEMK 73
Db 312 KRKEYEMK 319

RESULT 8
SUI1_SUITO STANDARD; PRT; 100 AA.
AC Q975S0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Protein translation factor SUI1 homolog.
GN ST0350.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=JCM 10545 / 7;
RC MEDLINE=21456156; PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyana A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
CC -!- SIMILARITY: BELONGS TO THE SUI1 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF000982; BAB65330.1; -;
DR HAMAP; MF 00604; -; 1.
DR InterPro; IPR005872; SUI1_arch_bact.
DR InterPro; IPR001950; TIF_SUI1.
DR Pfam; PF01253; SUI1; 1.
DR TIGRFAMs; TIGR01158; SUI1_rel; 1.
DR PROSITE; PS01118; SUI1_1; 1.
DR PROSITE; PS0296; SUI1_2; 1.
KW Protein biosynthesis; Translation regulation; Complete proteome.
SQ SEQUENCE 100 AA; 1185 MW; F0998DEALB633DB3 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 100;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 SDVELKK 28
Db 46 SDVELKK 52

RESULT 9
CB21_CARPI STANDARD; PRT; 111 AA.
ID CB21_CARPI
AC P38582;
DT 01-OCT-1994 (Rel. 30, Created)

01-OCT-1994 (Rel. 30, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Possible carnobacteriocin B2 immunity protein (Carnocin CP52 immunity
protein).
OS Carnobacterium piscicola.
DE Plasmid 61 kb.
OC Bacteria; Firmicutes; Lactobacillales; Carnobacteriaceae;
OC Carnobacterium.
OX NCBI_TaxID=2751;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LV17B;
RX MEDLINE=94216339; PubMed=8163526;
RA Quadri L.E.N., Sailer M., Roy K.L., Vederas J.C., Stiles M.E.;
RA "Chemical and genetic characterization of bacteriocins produced by
RT Carnobacterium piscicola LV17B.";
RT J. Biol. Chem. 269:12204-12211(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CP5;
RA Herbin S., Lebrihi A., Lefebvre G.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COULD IMPART IMMUNITY TO CARNOBACTERIOCIN B2 TO
CC NATURALLY SENSITIVE HOST STRAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L29059; AA72432.1; -;
DR EMBL; L47121; AAB81311.1; -;
DR EMBL; U76763; AAB18990.1; -;
DR PIR; D53589; D53589.
KW Bacteriocin immunity; Plasmid.
SQ SEQUENCE 111 AA; 12666 MW; 8387EAC50444BA12 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 NEFLSKL 91
Db 24 NEFLSKL 30

RESULT 10
RL7_ANASP STANDARD; PRT; 129 AA.
ID RL7_ANASP
AC Q8YLJ5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L7/L12.
GN RPL1 OR RPL12 OR ALR5303.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.F., Kuritz T., Sasamoto S.,
RA Watanabe A., Iritiguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001)
CC -!- FUNCTION: Seems to be the binding site for several of the factors

involved in protein synthesis and appears to be essential for accurate translation (By similarity).
-!- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.

This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: AP003599; BAB77002.1; -.
PIR: AG2468; AG2468.
DR HAMAP; MF_00368; -; 1.
DR InterPro; IPR000206; Ribosomal_L12.
DR Pfam; PF00542; Ribosomal_L12; 1.
DR ProDom; PD001326; Ribosomal_L12; 1.
DR TIGRFAMs; TIGR00855; L12; 1.
DR Ribosomal protein; Complete proteome.
KW RIBOSOMAL PROTEIN; BY SIMILARITY.
FT INIT MET 0
KW SEQUENCE 129 AA; 13326 MW; F68F001CA95847BC CRC64;
SQ
Query Match 2.6%; Score 7; DB 1; Length 129;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 237 AFGAAAA 243
D 46 AFGAAAA 52
D
RESULT 11
GLB2_MORMR STANDARD; PRT; 149 AA.
ID GLB2_MORMR
AC 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Globin II.
OS Mordacia mordax (Southern hemisphere lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Mordacia.
OX NCBI_TaxID=7755;
RN [1]
RP SEQUENCE
RX MEDLINE=91248417; PubMed=2039605;
RA Hombrados I., Vidal Y., Rodewald K., Braunitzer G., Neuzil E.;
RT "The primary structure of the hemoglobins of a southern hemisphere lamprey (Mordacia mordax, Cyclostomata).";
RL Biochem. Hoppe-Seyler 372:49-56(1991).
CC -!- SUBUNIT: Monomer
CC -!- SIMILARITY: Belongs to the globin family.
DR PIR; S13459; S13459.
DR HSSP; P02208; 2LHB.
DR InterPro; IPR000971; Globin.
DR Pfam; PF00042; globin; 1.
DR PROSITE; PS01033; GLOBIN; 1.
KW Heme; Oxygen transport; Transport.
FT METAL 73
FT METAL 73
FT METAL 105
FT METAL 105
FT METAL 105
SQ SEQUENCE 149 AA; 16490 MW; 160DCB60B84E56A CRC64;
Query Match 2.6%; Score 7; DB 1; Length 149;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 55 RAAWDIV 61
D 20 RAAWDIV 26
D

RESULT 12
TONB_YEREN STANDARD; PRT; 255 AA.
ID TONB_YEREN
AC Q05740;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TonB protein.
GN TONB.
OS Yersinia enterocolitica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=630;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93204889; PubMed=8384290;
RA Koebnik R., Baeumler A.J., Heesemann J., Braun V., Hantke K.;
RT "The TonB protein of Yersinia enterocolitica and its interactions with TonB-box proteins.";
RL Mol. Genet. 237:152-160(1993).
CC -!- FUNCTION: INTERACTS WITH OUTER MEMBRANE RECEPTOR PROTEINS THAT CARRY OUT HIGH-AFFINITY BINDING AND ENERGY DEPENDENT UPTAKE INTO THE PERIPLASMIC SPACE OF SPECIFIC SUBSTRATES. IT COULD ACT TO TRANSDUCE ENERGY FROM THE CYTOPLASMIC MEMBRANE TO SPECIFIC ENERGY-REQUIRING PROCESSES IN THE OUTER MEMBRANE, RESULTING IN THE RELEASE INTO THE PERIPLASM OF LIGANDS BOUND BY THESE OUTER MEMBRANE PROTEINS (BY SIMILARITY).
CC -!- SUBUNIT: HOMODIMER. FORMS A COMPLEX WITH THE ACCESSORY PROTEINS EXBB AND EXBD (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: PERIPLASMIC. ANCHORED TO THE CYTOPLASMIC MEMBRANE VIA ITS N-TERMINAL SIGNAL-LIKE SEQUENCE, SPANS THE PERIPLASM.
CC -!- SIMILARITY: BELONGS TO THE TONB FAMILY.

This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: X67332; CAA47747.1; -.
DR PIR; S30290; S30290.
DR HSSP; P94739; 1IHR.
DR InterPro; IPR003538; TonB.
DR Pfam; PF03544; TonB; 1.
DR PRINTS; PR01374; TONBPROTEIN.
DR TIGRFAMs; TIGR01352; tonB_Cterm; 1.
KW Transport; Protein transport; Inner membrane; Periplasmic; Transmembrane; Signal-anchor; Repeat.
FT DOMAIN 1 10
FT DOMAIN 11 33
FT TRANSMEM 34 255
FT DOMAIN 72 83
FT DOMAIN 94 105
FT DOMAIN 107 112
FT SEQUENCE 255 AA; 27861 MW; 0EA7285FDEBD988E CRC64;
SQ
Query Match 2.6%; Score 7; DB 1; Length 255;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 231 FAAPQPA 237
D 59 FAAPQPA 65
D
RESULT 13
MSX1_MOUSE STANDARD; PRT; 293 AA.
ID MSX1_MOUSE

AC P13297;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Homeobox protein MSX-1 (Hox-7) (Hox-7.1).
GN MSX1 OR HOX7 OR HOX7.1.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=920317197; PubMed=1602128;
RA Monaghan A.P., Davidson D.R., Sime C., Graham E., Baldock R.,
RA Bhattacharya S.S., Hill R.E.;
RT "The Msh-like homeobox genes define domains in the developing
RT vertebrate eye.";
RL Development 112:1053-1061(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89211899; PubMed=2565278;
RA Hill R.E., Jones P.F., Rees A.R., Sime C.M., Justice M.J.,
RA Copeland N.G., Jenkins N.A., Graham E., Davidson D.R.;
RT "A new family of mouse homeo box-containing genes: molecular
RT structure, chromosomal location, and developmental expression of
RT Hox-7.1.";
RL Genes Dev. 3:26-37(1989).
RN [3]
RP SEQUENCE OF 152-249 FROM N.A.
RX MEDLINE=89231641; PubMed=2565810;
RA Robert B., Sassoon D., Jacq B., Gehring W.J., Buckingham M.;
RT "Hox-7, a mouse homeobox gene with a novel pattern of expression
RT during embryogenesis.";
RL EMBO J. 8:91-100(1989).
RN [4]
RP SEQUENCE OF 165-225 FROM N.A.
RX MEDLINE=91200674; PubMed=1673109;
RA Holland P.W.H.;
RT "Cloning and evolutionary analysis of msh-like homeobox genes from
RT mouse, zebrafish and ascidian.";
RL Gene 98:253-257(1991).
RN [5]
RP FUNCTION.
RX MEDLINE=91319208; PubMed=1677742;
RA Davidson D.R., Crawley A., Hill R.E., Tickle C.;
RT "Position-dependent expression of two related homeobox genes in
RT developing vertebrate limbs.";
RL Nature 352:429-431(1991).
RN [6]
RP FUNCTION.
RX MEDLINE=95124344; PubMed=7823952;
RA Catron K.M., Zhang H., Marshall S.C., Inostroza J.A., Wilson J.M.,
RA Abate C.;
RT "Transcriptional repression by Max-1 does not require homeodomain
RT DNA-binding sites.";
RL Mol. Cell. Biol. 15:861-871(1995).
CC -!- FUNCTION: Acts as a transcriptional repressor. May play a role in
CC limb-pattern formation.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE MSH HOMEBOX FAMILY.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 246.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X59251; CAA41944.1; ALT_INIT.
DR

DR EMBL; X14759; CAA32871.1; ALT_FRAME.
DR EMBL; X14457; CAA32626.1; -.
DR EMBL; M38575; CAA37823.1; -.
DR PDB; 1IG7; 23-APR-01.
DR TRANSFAC; T02072; -.
DR MGD; MGI:97168; Msx1.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
KW Transcription regulation; Repressor; Homeobox; DNA-binding;
KW Developmental protein; Nuclear protein; 3D-structure.
FT DNA BIND 166 225 HOMEBOX.
SQ SEQUENCE 293 AA; 30846 MW; 27DC3B1B28693D8 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 293;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 AFGAAAA 243
Db 26 AFGAAAA 32
|||||
|

RESULT 14
JUND_CHICK
ID -JUND_CHICK STANDARD; PRT; 323 AA.
AC P27921;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Transcription factor jund.
GN JUND.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92019832; PubMed=1923529;
RA Hartl M., Hutchins J.T., Vogt P.K.;
RT "The chicken jund gene and its product.";
RL Oncogene 6:1623-1631(1991).
CC -!- SUBUNIT: Birds DNA as a dimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the bZIP family, Jun subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X60063; CAA42665.1; -.
DR PIR; S20099; S20099.
DR HSSP; P05412; LF0S.
DR TRANSFAC; T02196; -.
DR InterPro; IPR005643; JNK.
DR InterPro; IPR002112; Leuzip_Jun.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF00170; bZIP; 1.
DR Pfam; PF03957; Jun; 1.
DR PRINTS; PR00043; LEUZIPPJUN.
DR SMART; SM00338; BRIZ; 1.
DR PROSITE; PS50217; BZIP; 1.

DR PROSITE; PS00036; BZIP_BASIC; 1.
KW Transcription regulation; DNA-binding; Activator; Nuclear protein.
FT DOMAIN 59 67 POLY-ALA.
FT DOMAIN 155 166 POLY-GLY.
FT DNA_BIND 242 286 BASIC MOTIF.
FT DOMAIN 270 298 LEUCINE-ZIPPER.
SQ SEQUENCE 323 AA; 33205 MW; A7F6D21A97DBB676 CRC64;

Query Match 2.6%; Score 7; DB 1; Length 323;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 237 AFGAAAA 243
|||||
Db 56 AFGAAA 62

RESULT 15
GAS1_HUMAN STANDARD; PRT; 345 AA.
AC P54826;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Growth-arrest-specific protein 1 precursor (GAS-1).
GN GAS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A., AND FUNCTION.
RC TISSUE=Liver;
RX MEDLINE=94173926; PubMed=8127893;
RA del Sal G., Collavin L., Ruaro M.E., Edomi P., Saccone S., Valle G.D.,
RA Schneider C.;
RT "Structure, function, and chromosome mapping of the
RT growth-suppressing human homologue of the murine gas1 gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:1848-1852(1994).
CC !- FUNCTION: Specific growth arrest protein involved in growth
CC suppression. Blocks entry to S phase. Prevents cycling of
CC normal and transformed cells.
CC !- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
CC similarity).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L13698; AAA72368.1; -
CC F1R; A53138; A53138.
CC Genew; HGNC:4165; GAS1.
CC MIN; 139185; -
CC DR GO; GO:0007050; P:cell cycle arrest; TAS.
CC DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
KW Growth arrest; GPI-anchor; Signal; Glycoprotein.
FT SIGNAL 1 39 POTENTIAL.
FT CHAIN 40 318 GROWTH-ARREST-SPECIFIC PROTEIN 1
FT PROPEP 319 345 REMOVED IN MATURE FORM (POTENTIAL).
FT DOMAIN 87 91 POLY-ALA.
FT DOMAIN 285 288 POLY-ASP.
FT DOMAIN 319 322 POLY-GLY.
FT DOMAIN 337 341 POLY-LEU.
FT LIPID 318 318 GPI-ANCHOR (POTENTIAL).
FT CARBOHYD 117 117 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 345 AA; 35721 MW; 2AAD50F1D3632F9D CRC64;

Query Match 2.6%; Score 7; DB 1; Length 345;
Best Local Similarity 100.0%; Pred. No. 34;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 237 AFGAAAA 243
|||||
Db 84 AFGAAAA 90

Search completed: January 2, 2004, 16:45:36
Job time : 18 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 16:42:55 ; Search time 44 Seconds
(without alignments)
1577.640 Million cell updates/sec

Title: US-10-049-742-11

Perfect score: 269

Sequence: 1 MAGVPEDELNPFHVLGVEAT.....VPKGEAKPKRKVKRRPFQR 269

Scoring table: OLLIGO

Gapop 60.0 , Gapext 60.0

Searched: 830525 segs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL 23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertibrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	269	100.0	412	Q95N59	Q95N59 homo sapien
2	213	79.2	213	Q95T63	Q95T63 homo sapien
3	191	71.0	389	Q96AM4	Q96AM4 homo sapien
4	155	57.6	191	Q9C062	Q9C062 homo sapien
5	125	46.5	678	Q8BLF3	Q8BLF3 mus musculus
6	125	46.5	701	Q925G7	Q925G7 rattus norv
7	125	46.5	703	Q921R4	Q921R4 mus musculus
8	125	46.5	703	Q9CYE7	Q9CYE7 mus musculus
9	95	35.3	699	Q95J56	Q95J56 bos taurus
10	90	33.5	283	Q92EY9	Q92EY9 bovine vira
11	87	32.3	334	Q9E2W7	Q9E2W7 bovine vira
12	87	32.3	411	Q9E2W6	Q9E2W6 bovine vira
13	85	31.6	239	Q85797	Q85797 mucosal dis
14	81	30.1	3988	Q91WA7	Q91WA7 mucosal dis
15	81	30.1	3989	Q9PYB2	Q9PYB2 pestivirus
16	81	30.1	4040	Q91FH8	Q91FH8 mucosal dis

17	78	29.0	659	6	Q9BGH5	Q9BGH5 bos taurus
18	74	27.5	250	12	Q9E2W1	Q9E2W1 bovine vira
19	67	24.9	381	12	Q65787	Q65787 mucosal dis
20	67	24.9	703	11	Q8BUU3	Q8BUU3 mus musculus
21	64	23.8	309	12	Q65451	Q65451 bovine vira
22	61	22.7	371	12	Q9E2W3	Q9E2W3 bovine vira
23	53	19.7	237	12	Q65798	Q65798 mucosal dis
24	50	18.6	409	12	Q9E2V8	Q9E2V8 bovine vira
25	49	18.2	291	12	Q65450	Q65450 bovine vira
26	46	17.1	560	11	Q8BKF3	Q8BKF3 mus musculus
27	44	16.4	417	12	Q9E2W0	Q9E2W0 bovine vira
28	36	13.4	419	12	Q9E2W2	Q9E2W2 bovine vira
29	14	5.2	648	5	Q961F2	Q961F2 drosophila
30	14	5.2	970	5	Q9VN28	Q9VN28 drosophila
31	9	3.3	149	10	Q9C7C7	Q9C7C7 arabidopsis
32	9	3.3	262	10	Q9LH49	Q9LH49 arabidopsis
33	9	3.3	515	17	Q96YU3	Q96YU3 sulfolobus
34	8	3.0	163	4	Q96WF3	Q96WF3 homo sapien
35	8	3.0	182	11	Q9CZV4	Q9CZV4 mus musculus
36	8	3.0	182	11	Q9WUL8	Q9WUL8 mus musculus
37	8	3.0	186	16	Q97P65	Q97P65 streptococc
38	8	3.0	186	16	Q8CYE4	Q8CYE4 streptococc
39	8	3.0	289	5	Q8SS98	Q8SS98 encephalito
40	8	3.0	309	5	O15568	O15568 trichomonas
41	8	3.0	309	5	O15569	O15569 trichomonas
42	8	3.0	348	8	Q9GON7	Q9GON7 arapauma gi
43	8	3.0	349	8	Q9G7U7	Q9G7U7 hypochilus
44	8	3.0	484	5	Q9Y1W0	Q9Y1W0 dictyosteli
45	8	3.0	587	2	Q8VQP7	Q8VQP7 burkholderi

ALIGNMENTS

RESULT 1

Q96N59 PRELIMINARY; PRT; 412 AA.

AC Q96N59; (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ31383.
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.

RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Magatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA Isogai T.;

RT "NEDO human cDNA sequencing project.";

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK055945; BAB71050.1; --

DR InterPro; IPR001623; DnaJ_N.

DR InterPro; IPR003095; Hsp_DnaJ.

DR Pfam; PF00226; DnaJ; 1.

DR PRINTS; PR00625; DNAJPROTEIN.

DR SMART; SM00271; DnaJ; 1.

DR PROSITE; PS00076; DnaJ_2; 1.

SW Hypothetical protein.

SQ SEQUENCE 412 AA; 46931 MW; B51386515456C378 CRC64;

Query Match 100.0%; Score 269; DB 4; Length 412;

Best Local Similarity 100.0%; Pred. No. 4e-272;

Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MAGVPEDELNPFHVLGVEATSDVELKKAYRQLAVMVHPDKNHHPRAEAFKVLRAAWDI 60

```
Db 144 MAGVPEDELNPHVLGVETATSDVLEKKAYRQLAVVHPDKNHPRAEEAFKVLRAAWDI 203
QY 61 VSAEKREYEMKRAENELSRVNEFLSKLQDDLKEAMNTMWCSCQCKHRRFEMDREP 120
Db 204 VSAEKREYEMKRAENELSRVNEFLSKLQDDLKEAMNTMWCSCQCKHRRFEMDREP 263
QY 121 KSARYCAECNRLHPAEGDFWAESSMLGLKITYFALMDGKVVYDITWAGCQVRVIGISPDTH 180
Db 264 KSARYCAECNRLHPAEGDFWAESSMLGLKITYFALMDGKVVYDITWAGCQVRVIGISPDTH 323
QY 181 RVPYHISFGSRIPGTRGRQATPDAPPADLQDLSRIFQVPPGQMPNGNFFAAPPQAPGA 240
Db 324 RVPYHISFGSRIPGTRGRQATPDAPPADLQDLSRIFQVPPGQMPNGNFFAAPPQAPGA 383
QY 241 AAASXPNSTVPKGEAKPKRKKVRRPFQR 269
Db 384 AAASXPNSTVPKGEAKPKRKKVRRPFQR 412

RESULT 2
Q96T63 PRELIMINARY; PRT; 213 AA.
AC Q96T63;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Dopamine receptor interacting protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RX MEDLINE=21231375; PubMed=11331877;
RA Bermak J.C., Li M., Bullock C.M., Zhou Q.-Y.;
RT "Regulation of transport of the dopamine D1 receptor by a new
  membrane-associated ER protein.";
RL Nat. Cell Biol. 3:492-498(2001).
DR EMBL: AF351784; AAK56241.1; -.
KW Receptor.
FT NON_TER
SQ SEQUENCE 213 AA; 24206 MW; E7AF40EAD9086613 CRC64;

Query Match 79.2%; Score 213; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 6.9e-214;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 AWDIVSNAEKREYEMKRAENELSRVNEFLSKLQDDLKEAMNTMWCSCQCKHRRFEM 116
Db 1 AWDIVSNAEKREYEMKRAENELSRVNEFLSKLQDDLKEAMNTMWCSCQCKHRRFEM 60
QY 117 DRPKSARYCAECNRLHPAEGDFWAESSMLGLKITYFALMDGKVVYDITWAGCQVRVIGIS 176
Db 61 DRPKSARYCAECNRLHPAEGDFWAESSMLGLKITYFALMDGKVVYDITWAGCQVRVIGIS 120
QY 177 PDTHRVYPYHISFGSRIPGTRGRQATPDAPPADLQDLSRIFQVPPGQMPNGNFFAAPPQ 236
Db 121 PDTHRVYPYHISFGSRIPGTRGRQATPDAPPADLQDLSRIFQVPPGQMPNGNFFAAPPQ 180
QY 237 APGAAASXPNSTVPKGEAKPKRKKVRRPFQR 269
Db 181 APGAAASXPNSTVPKGEAKPKRKKVRRPFQR 213

RESULT 3
Q96AM4 PRELIMINARY; PRT; 389 AA.
AC Q96AM4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RC Tissue=Placenta;
RA Straubeberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC016941; AAH16941.1; -.
DR InterPro: IPR003034; SAP.
DR Pfam: PF02037; SAP; 1.
DR SMART: SM00513; SAP; 1.
DR Hypothetical protein.
SQ SEQUENCE 389 AA; 43745 MW; 3F8D3C3868031E66 CRC64;

Query Match 71.0%; Score 191; DB 4; Length 389;
Best Local Similarity 100.0%; Pred. No. 1.1e-190;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 MKRMAENELSRVNEFLSKLQDDLKEAMNTMWCSCQCKHRRFEMDREPKSARYCAECNR 131
Db 1 MKRMAENELSRVNEFLSKLQDDLKEAMNTMWCSCQCKHRRFEMDREPKSARYCAECNR 60
QY 132 LHPAEGDFWAESSMLGLKITYFALMDGKVVYDITWAGCQVRVIGISPDTHRVYPYHISFGSR 191
Db 61 LHPAEGDFWAESSMLGLKITYFALMDGKVVYDITWAGCQVRVIGISPDTHRVYPYHISFGSR 120
QY 192 IPGTRGRQATPDAPPADLQDLSRIFQVPPGQMPNGNFFAAPPQAPGAAASXPNSTVP 251
Db 121 IPGTRGRQATPDAPPADLQDLSRIFQVPPGQMPNGNFFAAPPQAPGAAASXPNSTVP 180
QY 252 KGEAKPKRKK 262
Db 181 KGEAKPKRKK 191

RESULT 4
Q9C062 PRELIMINARY; PRT; 191 AA.
AC Q9C062;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE LYST-interacting protein Lip6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RA Tchernev V.T., Mansfield T.A., Giot L., Kumar A.M., Nandabalan K.,
  Li Y., Mishra V.S., Dettler J.C., Rothberg J.M., Wallace M.R.,
  Southwick F.S., Kingmore S.F.;
RT "Interactions of the Chediak-Higashi lysosomal-trafficking regulator
  protein with SNARE complex and signal transduction proteins.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF141342; AAG49445.1; -.
SQ SEQUENCE 191 AA; 21675 MW; C9AE7DAE2775B101 CRC64;

Query Match 57.6%; Score 155; DB 4; Length 191;
Best Local Similarity 100.0%; Pred. No. 2.4e-153;
Matches 155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 MKRMAENELSRVNEFLSKLQDDLKEAMNTMWCSCQCKHRRFEMDREPKSARYCAECNR 131
Db 1 MKRMAENELSRVNEFLSKLQDDLKEAMNTMWCSCQCKHRRFEMDREPKSARYCAECNR 60
QY 132 LHPAEGDFWAESSMLGLKITYFALMDGKVVYDITWAGCQVRVIGISPDTHRVYPYHISFGSR 191
Db 61 LHPAEGDFWAESSMLGLKITYFALMDGKVVYDITWAGCQVRVIGISPDTHRVYPYHISFGSR 120
QY 192 IPGTRGRQATPDAPPADLQDLSRIFQVPPGQMP 226
```

Db 121 IPFTRGRQATDAPPADLQDFLSRIFQVPEQMP 155

RESULT 5

Q8BLUF3 Q925G7 PRELIMINARY; PRT; 678 AA.

AC Q8BLF3;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Dopamine receptor interacting protein homolog.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OC NCBI_TaxID=10090;

RP [1]

RN SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Brain;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RL 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

DR ENBL; AK045345; BAC32320.1; --

SQ SEQUENCE 678 AA; 76463 MW; 005EPEFF7F9AE5BF CRC64;

Query Match 46.5%; Score 125; DB 11; Length 678;

Best Local Similarity 100.0%; Pred.No. 1.6e-121; Gaps 0;

Matches 125; Conservative 0; Mismatches 0; Indels 0;

Qy 67 RREYEMKRWAEENELSRVNEFLSKLQDDLKEAMNTMCSRQCGKHRRFEMDREPKSARYC 126

Db 501 RREYEMKRWAEENELSRVNEFLSKLQDDLKEAMNTMCSRQCGKHRRFEMDREPKSARYC 560

Qy 127 AECNRLHPAEGDFWAESSMLGLKITYPALMDGKVYDITWAGQQRVIGISPDTHRVPYHI 186

Db 561 AECNRLHPAEGDFWAESSMLGLKITYPALMDGKVYDITWAGQQRVIGISPDTHRVPYHI 620

Qy 187 SFGSR 191

Db 621 SFGSR 625

RESULT 6

Q925G7 Q925G7 PRELIMINARY; PRT; 701 AA.

AC Q925G7;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Dopamine receptor interacting protein.

GN DRIP78.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OC NCBI_TaxID=10116;

RP [1]

RN SEQUENCE FROM N.A.

RC STRAIN=Sprague-Dawley;

RX MEDLINE=21231375; PubMed=11331877;

RA Bermak J.C., Li M., Bullock C.M., Zhou Q.-Y.;

RT "Regulation of transport of the dopamine D1 receptor by a new

RL membrane-associated ER protein.";

RL Nat. Cell Biol. 3:492-498(2001).

DR ENBL; AF351783; AAK56240.1; --

DR InterPro; IPR001623; DnaJ N.

DR InterPro; IPR003095; Hsp_DnaJ.

DR Pfam; PF00226; DnaJ_1.

DR PRINTS; PR00625; DNAJPROTEIN.

DR SMART; SM00271; DnaJ_1.

DR PROSITE; PS00076; DnaJ_2; 1.

GN 5730551F12RIK
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner T., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King C., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshak-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR ENBL; AK017830; BAE30962.1; -
DR ENBL; AK045445; BAC32372.1; -
DR HSP; P25685; IHDJ.
DR MGD; MGI:1921580; 5730551F12RIK.
DR InterPro; IPR001623; DnaJ N.
DR InterPro; IPR003095; Hsp_DnaJ.
DR Pfam; PF00226; DnaJ; 1.
DR PRINTS; PR00625; DnaJPROTEIN.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS50076; DnaJ 2; 1.
SQ SEQUENCE 703 AA; 78919 MW; B320DB4FB06345FE CRC64;
Query Match 46.5%; Score 125; DB 11; Length 703;
Best Local Similarity 100.0%; Pred. No. 1.7e-121;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 67 RKEYEMKRAENELSRVNEFLSKLQDLKEAMNTMWCSCQKRRFFEMDRPKSARYC 126
Db 501 RKEYEMKRAENELSRVNEFLSKLQDLKEAMNTMWCSCQKRRFFEMDRPKSARYC 560
QY 127 AECNRLHPAEGDFWAESSMLGLKITYFALMDGKVYDITWAGQCVGISPDTHRPVYHI 186
Db 561 AECNRLHPAEGDFWAESSMLGLKITYFALMDGKVYDITWAGQCVGISPDTHRPVYHI 620
QY 187 SFGRS 191
Db 621 SFGRS 625
RESULT 9
Q95J56 PRELIMINARY; PRT; 699 AA.
AC Q95J56;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE J-domain protein Jiv.
OS Bos taurus (Bovine)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21424530; PubMed=11533209;
RA Rinck G., Birghan C., Harada T., Meyers G., Thiel H.J., Tautz N.;
RT "A cellular J-domain protein modulates polyprotein processing and
RT cytopathogenicity of a pestivirus."
RL J. Virol. 75:9470-9482(2001).
DR EMBL; AY027882; AAK38651.1; -
DR EMBL; AY027881; AAK38650.1; -
DR InterPro; IPR001623; DnaJ N.
DR InterPro; IPR003095; Hsp_DnaJ.
DR Pfam; PF00226; DnaJ; 1.
DR PRINTS; PR00625; DnaJPROTEIN.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS50076; DnaJ 2; 1.
SQ SEQUENCE 699 AA; 78178 MW; 370F62297595F516 CRC64;
Query Match 35.3%; Score 95; DB 6; Length 699;
Best Local Similarity 100.0%; Pred. No. 3.7e-90;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 97 EAMNTMWCSCQKRRFFEMDRPKSARYCAECNRLHPAEGDFWAESSMLGLKITYFAL 156
Db 527 EAMNTMWCSCQKRRFFEMDRPKSARYCAECNRLHPAEGDFWAESSMLGLKITYFAL 586
QY 157 MDGKVYDITWAGQCVGISPDTHRPVYHISFGSR 191
Db 587 MDGKVYDITWAGQCVGISPDTHRPVYHISFGSR 621
RESULT 10
Q9E2V9 PRELIMINARY; PRT; 283 AA.
AC Q9E2V9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Polyprotein (Fragment).
OS Bovine viral diarrhoea virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Pestivirus.
OX NCBI_TaxID=136447;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BVDV2-5912C;
RX MEDLINE=20411455; PubMed=10954582;
RA Ridepath J.F., Neill J.D.;
RT "Detection and characterization of genetic recombination in cytopathic
RT type 2 bovine viral diarrhoea viruses."
RL J. Virol. 74:8771-8774(2000).
DR EMBL; AF268179; AAG13372.1; -
DR MEROPS; S31.001; -
DR NON_TER 1
DR NON_TER 283
SQ SEQUENCE 283 AA; 31949 MW; 440D47EAE8FB9585 CRC64;
Query Match 33.5%; Score 90; DB 12; Length 283;
Best Local Similarity 100.0%; Pred. No. 2.7e-85;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 97 EAMNTMWCSCQKRRFFEMDRPKSARYCAECNRLHPAEGDFWAESSMLGLKITYFAL 156
Db 70 EAMNTMWCSCQKRRFFEMDRPKSARYCAECNRLHPAEGDFWAESSMLGLKITYFAL 129
QY 157 MDGKVYDITWAGQCVGISPDTHRPVYHI 186

Db 130 MDGKVYDITEWAGCQVIGISPDTHRVPHI 159

Best Local Similarity 100.0%; Pred. No. 5.2e-82;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 11

Q9E2W7 PRELIMINARY; PRT; 334 AA.

ID Q9E2W7;
AC Q9E2W7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Polypeptide (Fragment)
OS Bovine viral diarrhoea virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Pestivirus.
OX NCBI_TaxID=136447;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BVDV2-296C;
RX MEDLINE=20411455; PubMed=10954582;
RA Ridpath J.F., Neill J.D.;
RT "Detection and characterization of genetic recombination in cytopathic
RT type 2 bovine viral diarrhoea viruses.";
RL J. Virol. 74:8771-8774 (2000).
DR EMBL; AF268171; AAG13364.1; -.
DR MEROPS; S31.001; -.
DR InterPro; IPR000280; CDvir_endptseP80.
DR PRINTS; PR00729; CDVENDOPTASE.
FT NON_TER 1
FT NON_TER 334
SQ SEQUENCE 334 AA; 37370 MW; F09D29D13D305476 CRC64;

Query Match 32.3%; Score 87; DB 12; Length 334;
Best Local Similarity 100.0%; Pred. No. 4.3e-82;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 EAMNTMCSRCQGHRRFEMDRPKSARYCAECNRLHPAEEDGFWAESMLGLKITYPAL 156
Db 16 EAMNTMCSRCQGHRRFEMDRPKSARYCAECNRLHPAEEDGFWAESMLGLKITYPAL 75

QY 157 MDGKVYDITEWAGCQVIGISPDTHRV 183
Db 76 MDGKVYDITEWAGCQVIGISPDTHRV 102

RESULT 12

Q9E2W6 PRELIMINARY; PRT; 411 AA.

ID Q9E2W6;
AC Q9E2W6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Polypeptide (Fragment)
OS Bovine viral diarrhoea virus type 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Pestivirus.
OX NCBI_TaxID=136447;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BVDV2-296C;
RX MEDLINE=20411455; PubMed=10954582;
RA Ridpath J.F., Neill J.D.;
RT "Detection and characterization of genetic recombination in cytopathic
RT type 2 bovine viral diarrhoea viruses.";
RL J. Virol. 74:8771-8774 (2000).
DR EMBL; AF268172; AAG13365.1; -.
DR MEROPS; S31.001; -.
DR InterPro; IPR000280; CDvir_endptseP80.
DR PRINTS; PR00729; CDVENDOPTASE.
FT NON_TER 1
FT NON_TER 411
SQ SEQUENCE 411 AA; 45794 MW; 31F1F60CCA611AD1 CRC64;

Query Match 32.3%; Score 87; DB 12; Length 411;

Db 130 MDGKVYDITEWAGCQVIGISPDTHRVPHI 159

Best Local Similarity 100.0%; Pred. No. 5.2e-82;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 13

Q65797 PRELIMINARY; PRT; 239 AA.

ID Q65797;
AC Q65797;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Viral nonstructural protein p125 (Fragment)
GN P125.
OS Mucosal disease virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Pestivirus.
OX NCBI_TaxID=11099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Indiana;
RA Greiser-Wilke I.M., Fritzscheimer J., Haas L.;
RT "Development of acute vs. late onset mucosal disease is a consequence
RT of different pathogenic mechanisms.";
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Indiana;
RA Fritzscheimer J.;
RL Thesis (1996), Institute of Virology, Veterinary School Hannover.
DR EMBL; 254331; CAA91136.1; -.
KW Nonstructural protein.
FT NON_TER 1
FT CHAIN 1
FT CHAIN 127
FT NON_TER 239
FT NON_TER 239
SQ SEQUENCE 239 AA; 27402 MW; F98FF1F8A4421A6 CRC64;

Query Match 31.6%; Score 85; DB 12; Length 239;
Best Local Similarity 100.0%; Pred. No. 3.9e-80;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 MNTMWCSSCQGHRRFEMDRPKSARYCAECNRLHPAEEDGFWAESMLGLKITYPALMD 158
Db 22 MNTMWCSSCQGHRRFEMDRPKSARYCAECNRLHPAEEDGFWAESMLGLKITYPALMD 81

QY 159 GKVDITEWAGCQVIGISPDTHRV 183
Db 82 GKVDITEWAGCQVIGISPDTHRV 106

RESULT 14

Q9IWA7 PRELIMINARY; PRT; 3988 AA.

ID Q9IWA7;
AC Q9IWA7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Genome polyprotein.
OS Mucosal disease virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Pestivirus.
OX NCBI_TaxID=11099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=type 1;

RX MEDLINE=20473135; PubMed=11018279;
RA Vassiliev V.B., Denis R.O.;
RT "Bovine viral diarrhoea virus induced apoptosis correlates with
RT increased intracellular viral RNA accumulation.";
RL Virus Res. 69:95-107(2000).
DR EMBL; AJ133738; CAB91846.1; -
DR HSP; P27958; 1A1V.
DR MEROPS; S31.001; -
DR InterPro; IPR000280; CDvir_endptsep80.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR001005; Myb DNA binding.
DR InterPro; IPR001568; RNase T2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF00271; helicase C; 1.
DR Pfam; PF00998; Viral RdRP; 1.
DR PRINTS; PR00729; CDVENDOPTASE.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
DR PROSITE; PS00037; MYB 1; 1.
DR PROSITE; PS05057; RDRP POSITIVE; 1.
DR PROSITE; PS0521; RDRP VIRAL; 1.
DR PROSITE; PS0531; RNASE T2; 1.
DR PROSITE; PS0533; RNASE T2; 1.
KW ATP-binding; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
KW RNA-directed RNA polymerase; Transferase.
FT CHAIN 3988 AA; 449122 MW; 8DCCF24D64C04A6B CRC64;
SQ SEQUENCE 3988 AA; 449122 MW; 8DCCF24D64C04A6B CRC64;

Query Match 30.1%; Score 81; DB 12; Length 3988;
Best Local Similarity 100.0%; Pred. No. 7.3e-75;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 MCSRCQCKHRRFEMDRPKSARYCAECNRLHPAEEDGDFWAESSMLGLKITYFALMDGKVY 162
DB 1537 MCSRCQCKHRRFEMDRPKSARYCAECNRLHPAEEDGDFWAESSMLGLKITYFALMDGKVY 1596

QY 163 DITEWAGCQVRGISPDTTHRVP 183
DB 1597 DITEWAGCQVRGISPDTTHRVP 1617

RESULT 15
Q9PYB2 PRELIMINARY; PRT; 3989 AA.
AC Q9PYB2
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Genome polyprotein.
OS Pestivirus Giraffe-1.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Pestivirus.
OX NCBI_TaxID=155905;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Giraffe-1 H138;
RX MEDLINE=99420379; PubMed=10489341;
RA Becher P., Orlich M., Kosmidou A., Konig M., Baroth M., Thiel H.J.;
RT "Genetic diversity of pestiviruses: identification of novel groups and
RT implications for classification.";
RL Virology 262:64-71(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Giraffe-1 H138;
RX MEDLINE=21378881; PubMed=11485413;
RA Avalos-Ramirez R., Orlich M., Thiel H.J., Becher P.;
RT "Evidence for the presence of two novel pestivirus species.";
RL Virology 286:456-465(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Giraffe-1 H138;
RA Avalos-Ramirez R., Orlich M., Thiel H.-J., Becher P.;

RT "Complete genomic sequences of pestiviruses from giraffe and reindeer:
RT evidence for the presence of two novel species within the genus
RT pestivirus.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF144617; AAF02523.2; -
DR MEROPS; S31.001; -
DR InterPro; IPR000280; CDvir_endptsep80.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR001005; Myb DNA binding.
DR InterPro; IPR001568; RNase T2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF00271; helicase C; 1.
DR Pfam; PF00998; Viral RdRP; 1.
DR PRINTS; PR00729; CDVENDOPTASE.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
DR PROSITE; PS00037; MYB 1; 1.
DR PROSITE; PS05057; RDRP POSITIVE; 1.
DR PROSITE; PS0521; RDRP VIRAL; 1.
DR PROSITE; PS0531; RNASE T2; 1.
KW ATP-binding; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
KW RNA-directed RNA polymerase; Transferase.
FT CHAIN 3989 AA; 447358 MW; E8CC6E4897FFE01A CRC64;
SQ SEQUENCE 3989 AA; 447358 MW; E8CC6E4897FFE01A CRC64;

Query Match 30.1%; Score 81; DB 12; Length 3989;
Best Local Similarity 100.0%; Pred. No. 7.3e-75;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 RCQGHRRFEMDRPKSARYCAECNRLHPAEEDGDFWAESSMLGLKITYFALMDGKVYDIT 165
DB 1536 RCQGHRRFEMDRPKSARYCAECNRLHPAEEDGDFWAESSMLGLKITYFALMDGKVYDIT 1595

QY 166 EWAGCQVRGISPDTTHRVPYHI 186
DB 1596 EWAGCQVRGISPDTTHRVPYHI 1616

Search completed: January 2, 2004, 16:46:34
Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 16:44:39 ; Search time 21 Seconds
(without alignments)
541.982 Million cell updates/sec

Title: US-10-049-742-11
Perfect score: 269
Sequence: 1 MAGVPEDELNPPHVLGVEAT.....VPKGEAKPKRKKVRPFQR 269

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	3.0	173	4	US-09-252-991A-16948
2	8	3.0	215	4	US-09-252-991A-28157
3	8	3.0	295	2	US-08-504-459-14
4	8	3.0	297	2	US-08-504-459-12
5	8	3.0	299	2	US-08-504-459-10
6	8	3.0	351	4	US-09-252-991A-21647
7	8	3.0	511	4	US-09-328-352-6176
8	8	3.0	686	4	US-09-252-991A-20587
9	8	3.0	1096	4	US-09-252-991A-19328
10	8	3.0	1474	4	US-09-241-606-2
11	7	2.6	111	4	US-08-924-629C-48
12	7	2.6	111	4	US-08-924-629C-51
13	7	2.6	111	4	US-08-370-838-199
14	7	2.6	137	4	US-09-252-991A-20631
15	7	2.6	149	4	US-09-107-532A-5732
16	7	2.6	223	4	US-09-252-991A-32114
17	7	2.6	245	4	US-09-252-991A-30805
18	7	2.6	358	4	US-09-679-279-16
19	7	2.6	383	3	US-09-413-814-12
20	7	2.6	397	2	US-08-868-288A-5
21	7	2.6	397	3	US-09-235-373-5
22	7	2.6	397	3	US-09-388-993-5
23	7	2.6	402	4	US-09-252-991A-26529
24	7	2.6	420	1	US-07-700-526-1
25	7	2.6	420	5	PCT-US92-03132-1
26	7	2.6	424	4	US-09-252-991A-30209
27	7	2.6	426	3	US-08-561-083-48

ALIGNMENTS

RESULT 1
US-09-252-991A-16948
; Sequence 16948, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,789
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16948
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16948

Query Match 3.0%; Score 8; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 PAPGAAAA 243
Db 5 PAPGAAAA 12

RESULT 2
US-09-252-991A-28157
; Sequence 28157, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28157
; LENGTH: 215
; TYPE: PRT


```
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28157

Query Match          3.0%; Score 8; DB 4; Length 215;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 236 PAPGAAAA 243
DB 136 PAPGAAAA 143
|||||

RESULT 3
US-08-504-459-14
; Sequence 14, Application US/08504459
; Patent No. 5922563
; GENERAL INFORMATION:
; APPLICANT: Alderete, John F.
; TITLE OF INVENTION: Adhesin Genes and Proteins Involved in
; TELECOMMUNICATION INFORMATION:
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: US
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/504,459
; FILING DATE: Concurrently herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mark B.
; REGISTRATION NUMBER: 37,259
; REFERENCE/DOCKET NUMBER: UTSK:273/WIM
; TITLE OF INVENTION: Trichomonas Vaginalis Cytoadherence
; TELECOMMUNICATION INFORMATION:
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: US
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/504,459
; FILING DATE: Concurrently herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mark B.
; REGISTRATION NUMBER: 37,259
; REFERENCE/DOCKET NUMBER: UTSK:273/WIM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 295 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-504-459-14

Query Match          3.0%; Score 8; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 236 PAPGAAAA 243
DB 71 PAPGAAA 78
|||||

RESULT 4
US-08-504-459-12
; Sequence 12, Application US/08504459
; Patent No. 5922563
; GENERAL INFORMATION:
; APPLICANT: Alderete, John F.
; TITLE OF INVENTION: Adhesin Genes and Proteins Involved in
; TELECOMMUNICATION INFORMATION:
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: US
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/504,459
; FILING DATE: Concurrently herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mark B.
; REGISTRATION NUMBER: 37,259
; REFERENCE/DOCKET NUMBER: UTSK:273/WIM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 295 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-504-459-14

Query Match          3.0%; Score 8; DB 2; Length 297;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 236 PAPGAAAA 243
DB 73 PAPGAAA 80
|||||

RESULT 5
US-08-504-459-10
; Sequence 10, Application US/08504459
; Patent No. 5922563
; GENERAL INFORMATION:
; APPLICANT: Alderete, John F.
; TITLE OF INVENTION: Adhesin Genes and Proteins Involved in
; TELECOMMUNICATION INFORMATION:
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: US
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/504,459
; FILING DATE: Concurrently herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mark B.
; REGISTRATION NUMBER: 37,259
; REFERENCE/DOCKET NUMBER: UTSK:273/WIM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-504-459-12
```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 299 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-504-459-10

Query Match 3.0%; Score 8; DB 2; Length 299;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 PAPGAAA 243
| | | | |
Db 75 PAPGAAA 82

RESULT 6

US-09-252-991A-21647
; Sequence 21647, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21647
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21647

Query Match 3.0%; Score 8; DB 4; Length 351;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 PAPGAAA 243
| | | | |
Db 274 PAPGAAA 281

RESULT 7

US-09-328-352-6176
; Sequence 6176, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328.352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6176
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6176

Query Match 3.0%; Score 8; DB 4; Length 511;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EATASDVE 25
| | | | |
Db 57 EATASDVE 64

RESULT 8

US-09-252-991A-20587
; Sequence 20587, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20587
; LENGTH: 696
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (516),(549),(562),(599)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-20587

Query Match 3.0%; Score 8; DB 4; Length 696;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 RIPGTRGR 198
| | | | |
Db 650 RIPGTRGR 657

RESULT 9

US-09-252-991A-19328
; Sequence 19328, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19328
; LENGTH: 1096
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19328

Query Match 3.0%; Score 8; DB 4; Length 1096;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 APGAAAAS 244
| | | | |
Db 350 APGAAAAS 357

RESULT 10

US-09-241-606-2
; Sequence 2, Application US/09241606
; Patent No. 6472140
; GENERAL INFORMATION:
; APPLICANT: Tanzi, Rudolph E.

; APPLICANT: Kovacs, Dora
; APPLICANT: Saunders, Aleister J.
; TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for
; TITLE OF INVENTION: Alzheimer's Disease
; FILE REFERENCE: 0609.4460003
; CURRENT APPLICATION NUMBER: US/09/241.606
; CURRENT FILING DATE: 1999-02-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1474
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-241-606-2

Query Match 3.0%; Score 8; DB 4; Length 1474;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 KRKEYEMK 73
Db 312 KRKEYEMK 319

RESULT 11
US-08-924-629C-48
; Sequence 48, Application US/08924629C
; Patent No. 6403082
; GENERAL INFORMATION:
; APPLICANT: Stiles, Michael E.
; APPLICANT: Vederas, John C.
; APPLICANT: van Belkum, Marius J.
; APPLICANT: Worobo, Randy W.
; APPLICANT: Greer, G. Gordon
; APPLICANT: McMullen, Lynn M.
; APPLICANT: Leisner, Jorgen J.
; APPLICANT: Poon, Aision
; TITLE OF INVENTION: Franz, Charles M.A.P.
; FILE REFERENCE: 660.0005US
; CURRENT APPLICATION NUMBER: US/08/924.629C
; CURRENT FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/026,257
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 111
; TYPE: PRT
; ORGANISM: carnobacterium B2 operon;
US-08-924-629C-48

Query Match 2.6%; Score 7; DB 4; Length 111;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 NEFLSKL 91
Db 24 NEFLSKL 30

RESULT 12
US-08-924-629C-51
; Sequence 51, Application US/08924629C
; Patent No. 6403082
; GENERAL INFORMATION:
; APPLICANT: Stiles, Michael E.
; APPLICANT: Vederas, John C.
; APPLICANT: van Belkum, Marius J.
; APPLICANT: Worobo, Randy W.
; APPLICANT: Greer, G. Gordon

; APPLICANT: McMullen, Lynn M.
; APPLICANT: Leisner, Jorgen J.
; APPLICANT: Poon, Aision
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: No. 6403082elBacteriocins, Transport and Vector System and Method
; FILE REFERENCE: 660.0005US
; CURRENT APPLICATION NUMBER: US/08/924.629C
; CURRENT FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/026,257
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Carnobacterium B2 operon
US-08-924-629C-51

Query Match 2.6%; Score 7; DB 4; Length 111;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 NEFLSKL 91
Db 24 NEFLSKL 30

RESULT 13
US-09-370-838-199
; Sequence 199, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370.838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 199
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-370-838-199

Query Match 2.6%; Score 7; DB 4; Length 127;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 ELKKAYR 31
Db 21 ELKKAYR 27

RESULT 14
US-09-252-991A-20631
; Sequence 20631, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 20631
 ; LENGTH: 137
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-20631

Query Match 2.6%; Score 7; DB 4; Length 137;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 PAGAAA 242
 Db 65 PAGAAA 71

RESULT 15

US-09-107-532A-5732
 ; Sequence 5732, Application US/09107532A
 ; Patent No. 6583275
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 7310
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 ; STREET: 100 Beaver Street
 ; CITY: Waltham
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02354
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD-ROM ISO9660
 ; COMPUTER: PC
 ; OPERATING SYSTEM: <Unknown>
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/107,532A
 ; FILING DATE: 30-Jun-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/085,598
 ; FILING DATE: 14 May 1998
 ; APPLICATION NUMBER: 60/051571
 ; FILING DATE: July 2, 1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ariniello, Pamela Deneke
 ; REGISTRATION NUMBER: 40,489
 ; REFERENCE/DOCKET NUMBER: GTC-012
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (781)893-5007
 ; TELEFAX: (781)893-8277
 ; INFORMATION FOR SEQ ID NO: 5732:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 149 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: YES
 ; ORIGINAL SOURCE:
 ; ORGANISM: Enterococcus faecium
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (B) LOCATION 1...149
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 5732:
 US-09-107-532A-5732

Query Match 2.6%; Score 7; DB 4; Length 149;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 EEAFKVL 54
 Db 113 EEAFKVL 119

Search completed: January 2, 2004, 16:48:19
 Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: January 2, 2004, 16:45:34 ; Search time 32 Seconds
(without alignments)
1686.640 Million cell updates/sec

Title: US-10-049-742-11
Perfect score: 269
Sequence: 1 MAGVPEDELNPHVLGVEAT.....VPKGEAKPKRKVRFPQR 269

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 733937 seqs, 200641211 residues

Word size : 0

Total number of hits satisfying chosen parameters: 733937

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	269	100.0	412	12	US-10-094-749-2296 Sequence 2296, Ap
2	155	57.6	131	15	US-10-097-340-183 Sequence 183, App
3	8	3.0	163	12	US-10-094-749-2459 Sequence 2459, Ap
4	8	3.0	272	10	US-09-771-161A-149 Sequence 149, App
5	8	3.0	350	12	US-10-369-493-2402 Sequence 2402, Ap
6	8	3.0	1050	12	US-10-369-493-4141 Sequence 4141, Ap
7	8	3.0	1162	12	US-10-369-493-13727 Sequence 13727, A
8	8	3.0	1285	9	US-09-925-301-1394 Sequence 1394, Ap
9	8	3.0	1450	12	US-09-756-247-23 Sequence 23, Appl
10	8	3.0	1451	12	US-09-756-247-24 Sequence 24, Appl
11	8	3.0	1474	9	US-09-873-403-5 Sequence 5, Appl
12	8	3.0	1474	12	US-10-292-081A-10 Sequence 10, Appl
13	8	3.0	1474	12	US-10-292-081A-12 Sequence 12, Appl
14	8	3.0	1474	12	US-10-292-081A-13 Sequence 13, Appl
15	8	3.0	1474	12	US-10-292-081A-15 Sequence 15, Appl

Sequence 38, Appl
Sequence 2, Appl
Sequence 9, Appl
Sequence 3915, A
Sequence 48, Appl
Sequence 51, Appl
Sequence 47613, A
Sequence 14773, A
Sequence 10753, A
Sequence 199, App
Sequence 199, App
Sequence 199, App
Sequence 1163, Ap
Sequence 2436, A
Sequence 9074, Ap
Sequence 18, Appl
Sequence 10401, A
Sequence 19646, A
Sequence 15859, A
Sequence 15487, A
Sequence 22783, A
Sequence 8065, Ap
Sequence 178, App
Sequence 17, Appl
Sequence 48, Appl
Sequence 22765, A
Sequence 2468, Ap
Sequence 46, Appl
Sequence 210, App
Sequence 210, App
US-10-331-496A-38
US-10-052-817-2
US-10-232-081A-9
US-09-864-761-39915
US-09-883-343A-48
US-09-883-343A-51
US-09-864-761-47613
US-10-156-761-14773
US-10-156-761-10753
US-09-738-373-199
US-09-854-133-199
US-10-144-649A-199
US-09-764-868-1163
US-10-369-493-22436
US-10-156-761-9074
US-10-166-087-18
US-10-369-493-10401
US-10-369-493-19646
US-10-369-493-15859
US-10-369-493-15487
US-10-369-493-8065
US-09-919-039-178
US-09-770-509-17
US-09-765-272-48
US-10-369-493-22765
US-10-094-749-2468
US-10-120-637A-46
US-10-137-870-210
US-10-140-018-210

ALIGNMENTS

RESULT 1
US-10-094-749-2296
Sequence 2296, Application US/10094749
Publication No. US20030219741A1
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHIKO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2296
LENGTH: 412
TYPE: PRT
ORGANISM: Homo sapiens

US-10-094-749-2296

Query Match 100.0%; Score 269; DB 12; Length 412;
Best Local Similarity 100.0%; Pred. No. 3e-250;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGVPEDLNPFHVLGVATASDELKAYRQLAVMVHPDKNHHPRAEAEAFKVLRAWDI 60
DB 144 MAGVPEDLNPFHVLGVATASDELKAYRQLAVMVHPDKNHHPRAEAEAFKVLRAWDI 203

QY 61 VSNAEKKEVEMKMAENELSRVNEFLSKLQDDLEAMNTWMCSCQCKHRRFEMDREP 120
DB 204 VSNAEKKEVEMKMAENELSRVNEFLSKLQDDLEAMNTWMCSCQCKHRRFEMDREP 263

QY 121 KSARYCAECNRLHPAEEGDFWAESSMLGLKITFYALMDGKVYDITWAGCQRVIGISPDTH 190
DB 264 KSARYCAECNRLHPAEEGDFWAESSMLGLKITFYALMDGKVYDITWAGCQRVIGISPDTH 323

QY 181 RVPYHISFGSRIPGTRGQRATPDAPPADLQDFLSRIFQVPPGMPNGNFFAAPQAPGA 240
DB 324 RVPYHISFGSRIPGTRGQRATPDAPPADLQDFLSRIFQVPPGMPNGNFFAAPQAPGA 383

QY 241 AAASKPNSTVPKGEAKPKRKKVRRPFQR 269
DB 384 AAASKPNSTVPKGEAKPKRKKVRRPFQR 412

RESULT 2
US-10-097-340-183
; Sequence 183, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPARU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 183
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-183

Query Match 57.6%; Score 155; DB 15; Length 191;
Best Local Similarity 100.0%; Pred. No. 6.7e-141;
Matches 155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 MKRMAENELSRVNEFLSKLQDDLEAMNTWMCSCQCKHRRFEMDREPFSARYCAECNR 131
DB 1 MKRMAENELSRVNEFLSKLQDDLEAMNTWMCSCQCKHRRFEMDREPFSARYCAECNR 60

QY 132 LHPAEEGDFWAESSMLGLKITFYALMDGKVYDITWAGCQRVIGISPDTHRVPYHISFQSR 191
DB 61 LHPAEEGDFWAESSMLGLKITFYALMDGKVYDITWAGCQRVIGISPDTHRVPYHISFQSR 120

QY 192 IPGTRGQRATPDAPPADLQDFLSRIFQVPPGMP 226
DB 121 IPGTRGQRATPDAPPADLQDFLSRIFQVPPGMP 155

RESULT 3
US-10-094-749-2459
; Sequence 2459, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2459
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2459

Query Match 3.0%; Score 8; DB 12; Length 163;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 PAPGAAAA 243
DB 79 PAPGAAAA 86

RESULT 4
US-09-771-161A-149
; Sequence 149, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES

```
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771.161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 149
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-149

Query Match          3.0%; Score 8; DB 10; Length 272;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 PARGAAAA 243
Db 39 PARGAAAA 46

RESULT 5
US-10-369-493-2402
; Sequence 2402, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2402
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(350)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-2402

Query Match          3.0%; Score 8; DB 12; Length 350;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 SSMGLGLKI 151
Db 52 SSMGLGLKI 59

RESULT 6
US-10-369-493-4141
; Sequence 4141, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2402
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(350)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-2402

Query Match          3.0%; Score 8; DB 12; Length 350;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 SSMGLGLKI 151
Db 52 SSMGLGLKI 59

RESULT 7
US-10-369-493-13727
; Sequence 13727, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13727
; LENGTH: 1162
; TYPE: PRT
; ORGANISM: Pseudomonas fluorescens
US-10-369-493-13727

Query Match          3.0%; Score 8; DB 12; Length 1162;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 LODDLKEA 98
Db 317 LODDLKEA 324

RESULT 8
US-09-925-301-1394
; Sequence 1394, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
```

; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1394
; LENGTH: 1285
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1394

Query Match 3.0%; Score 8; DB 9; Length 1285;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 KRKEYEMK 73
Db 123 KRKEYEMK 130

RESULT 9

US-09-756-247-23
; Sequence 23, Application US/09756247
; Publication No. US20030180722A1
; GENERAL INFORMATION:
; APPLICANT: Godbole, Shubhada D
; APPLICANT: Boyle, Bryan J
; APPLICANT: Mize, Nancy K
; APPLICANT: Deng, Cenhua
; APPLICANT: Goodrich, Ryle
; APPLICANT: Arterburn, Matthew C
; APPLICANT: Zhou, Ping
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Yeung, George
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO ALPHA-2-MACROGLOBULIN-LIKE P
; TITLE OF INVENTION: AND POLYNUCLEOTIDES
; FILE REFERENCE: HYS-31CIP
; CURRENT APPLICATION NUMBER: US/09/756,247
; CURRENT FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/684,711
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 1450
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-756-247-23

Query Match 3.0%; Score 8; DB 12; Length 1450;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 KRKEYEMK 73
Db 299 KRKEYEMK 306

RESULT 10

US-09-756-247-24
; Sequence 24, Application US/09756247
; Publication No. US20030180722A1
; GENERAL INFORMATION:
; APPLICANT: Godbole, Shubhada D
; APPLICANT: Boyle, Bryan J

; APPLICANT: Mize, Nancy K
; APPLICANT: Deng, Cenhua
; APPLICANT: Goodrich, Ryle
; APPLICANT: Arterburn, Matthew C
; APPLICANT: Zhou, Ping
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Yeung, George
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO ALPHA-2-MACROGLOBULIN-LIKE P
; TITLE OF INVENTION: AND POLYNUCLEOTIDES
; FILE REFERENCE: HYS-31CIP
; CURRENT APPLICATION NUMBER: US/09/756,247
; CURRENT FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/684,711
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 1451
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-756-247-24

Query Match 3.0%; Score 8; DB 12; Length 1451;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 KRKEYEMK 73
Db 300 KRKEYEMK 307

RESULT 11

US-09-873-403-5
; Sequence 5, Application US/09873403
; Patent No. US20020028207A1
; GENERAL INFORMATION:
; APPLICANT: Srivastava, Pramod K
; TITLE OF INVENTION: COMPLEXES OF ALPHA (2) MACROGLOBULIN AND ANTIGENIC
; TITLE OF INVENTION: MOLECULES FOR IMMUNOTHERAPY
; FILE REFERENCE: 8449-178
; CURRENT APPLICATION NUMBER: US/09/873,403
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 09/625,139
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/209,266
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1474
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-873-403-5

Query Match 3.0%; Score 8; DB 9; Length 1474;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 KRKEYEMK 73
Db 312 KRKEYEMK 319


```
RESULT 12
US-10-292-081A-10
; Sequence 10, Application US/10292081A
; Publication No. US20030162202A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth David Becker
; APPLICANT: Gonul Velicelebi
; APPLICANT: Xin Wang
; APPLICANT: Randolph E. Tanzi
; APPLICANT: Lars Bertram
; APPLICANT: Aleister J. Saunders
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLOBULIN
; FILE REFERENCE: 37481-3323
; CURRENT APPLICATION NUMBER: US/10/292,081A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: 60/337434
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1474
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-081A-10
Query Match      3.0%; Score 8; DB 12; Length 1474;
Best Local Similarity 100.0%; Pred. No. 92;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      66 KRKEYEMK 73
Db      312 KRKEYEMK 319

RESULT 13
US-10-292-081A-12
; Sequence 12, Application US/10292081A
; Publication No. US20030162202A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth David Becker
; APPLICANT: Gonul Velicelebi
; APPLICANT: Xin Wang
; APPLICANT: Randolph E. Tanzi
; APPLICANT: Lars Bertram
; APPLICANT: Aleister J. Saunders
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLOBULIN
; FILE REFERENCE: 37481-3323
; CURRENT APPLICATION NUMBER: US/10/292,081A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: 60/337434
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1474
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-081A-12
Query Match      3.0%; Score 8; DB 12; Length 1474;
Best Local Similarity 100.0%; Pred. No. 92;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      66 KRKEYEMK 73
Db      312 KRKEYEMK 319

RESULT 14
US-10-292-081A-13
; Sequence 13, Application US/10292081A
; Publication No. US20030162202A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Kenneth David Becker
; APPLICANT: Gonul Velicelebi
; APPLICANT: Xin Wang
; APPLICANT: Randolph E. Tanzi
; APPLICANT: Lars Bertram
; APPLICANT: Aleister J. Saunders
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLOBULIN
; FILE REFERENCE: 37481-3323
; CURRENT APPLICATION NUMBER: US/10/292,081A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: 60/337434
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1474
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-081A-13
Query Match      3.0%; Score 8; DB 12; Length 1474;
Best Local Similarity 100.0%; Pred. No. 92;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      66 KRKEYEMK 73
Db      312 KRKEYEMK 319

RESULT 15
US-10-292-081A-15
; Sequence 15, Application US/10292081A
; Publication No. US20030162202A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth David Becker
; APPLICANT: Gonul Velicelebi
; APPLICANT: Xin Wang
; APPLICANT: Randolph E. Tanzi
; APPLICANT: Lars Bertram
; APPLICANT: Aleister J. Saunders
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLOBULIN
; FILE REFERENCE: 37481-3323
; CURRENT APPLICATION NUMBER: US/10/292,081A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: 60/337434
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 1474
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-081A-15
Query Match      3.0%; Score 8; DB 12; Length 1474;
Best Local Similarity 100.0%; Pred. No. 92;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      66 KRKEYEMK 73
Db      312 KRKEYEMK 319

Search completed: January 2, 2004, 16:48:44
Job time : 33 secs
```